Optimality Conditions and Numerical Methods for a Continuous Reformulation of Cardinality Constrained Optimization Problems

Dissertation

von

Max Bucher

am Fachbereich Mathematik
der Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doctor rerum naturalium (Dr. rer. nat.)

Erstgutachterin: Prof. Dr. Alexandra Schwartz
Zweitgutachter: Prof. Dr. Christian Kanzow

TECHNISCHE UNIVERSITÄT DARMSTADT

Darmstadt 2018

Bucher, Max : Optimality Conditions and Numerical Methods for a Continuous Reformulation of Cardinality Constrained Optimization Problems

Darmstadt, Technische Universität Darmstadt
Jahr der Veröffentlichung der Dissertation auf TUprints: 2018
Tag der mündlichen Prüfung: 01.06.2018

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Dr. Alexandra Schwartz for giving me the opportunity to work in her research group and suggesting this interesting research topic. She was always there when I needed help or a piece of advice. I am very grateful to Alexandra for her guidance and support from which I profited immensely over the past years.

I thank Prof. Dr. Anja Klein for acting as a co-supervisor of this thesis within the Graduate School Computational Engineering, and for giving me insights in the work of the Communications Engineering Lab at Technische Universität Darmstadt.

I am thankful to Prof. Dr. Christian Kanzow for acting as a referee of this thesis and for his helpful suggestions regarding local error bounds.

Furthermore, I thank Prof. Dr. Marc Pfetsch and PD Dr. Christian Stinner for their willingness to participate in my examination committee.

I thank Dr. Martin Branda and Dr. Michal Červinka for the fruitful collaboration on our publication concerning a Scholtes-type regularisation and sparse robust portfolio optimization.

I am grateful to Prof. Dr. Oliver Stein for his encouraging help in pursuing a master's degree in mathematics and starting my doctorate afterwards.

Many thanks go to my colleagues at the Research Group Optimization and the Graduate School Computational Engineering at Technische Universität Darmstadt for creating a great working atmosphere. I am grateful to Christian Schmitt for his valuable and reliable assistance with the computer systems. Furthermore, I am grateful to Jan Becker, Johanna Biehl, Tristan Gally, Philip Kolvenbach, Christopher Müller, Daniel Nowak, Anne-Therese Rauls, Michel Reiffert, Christopher Spannring and Anna Walter for proof-reading parts of this thesis and for their valuable suggestions.

I gratefully acknowledge the support of the Excellence Initiative of the German Federal and State Governments and the Graduate School of Computational Engineering at Technische Universität Darmstadt. Furthermore, I gratefully acknowledge the support of the Deutsche Forschungsgemeinshaft within the Collaborative Research Centre Transregio 154 (Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks) towards the end of my thesis.

Abstract

Nonlinear constrained optimization problems can be used to model practical and theoretical questions from a vast range of areas in science and industry. In this thesis, we consider a certain class of these problems: *cardinality constrained optimization problems.* The goal is to minimise a, possibly nonlinear, objective function subject to given constraints, which we also allow to be nonlinear. The *cardinality constraint* is of our particular interest. It restricts the maximum number of nonzero components of a feasible vector. Cardinality constraints play an important role in a range of applications such as portfolio optimization, compressed sensing as well as in logistics.
Since the cardinality constraint is given by a discontinuous function, theoretical results and methods from nonlinear optimization cannot be readily applied. In this thesis, we consider an approach by Burdakov et al. (2016): a reformulation of the cardinality constraint with a continuous auxiliary variable. This reformulation bears strong similarities to a mathematical program with complementarity constraints (MPCC). Similarly, it violates common constraint qualifications which ensure that first order optimality conditions hold at a local minimum. For this reason custom constraint qualifications and first order optimality condition were introduced by Burdakov et al. (2016) and Červinka et al. (2016).
In this thesis we follow the approach by Burdakov et. al. (2016). We introduce new second order optimality conditions for the reformulation which hold under custom constraint qualifications. These include a necessary second order optimality condition, a sufficient second order optimality condition and a result on the local uniqueness of stationary points. Additionally, we deduce counterparts of those results for the original cardinality constrained problem. Moreover, the existence of a local error bound for the reformulation is shown. These optimality conditions as well as the result on the existence of a local error bound are subsequently used for the convergence theories of numerical methods.
In this thesis, we furthermore place emphasis on numerical methods. We proof exactness of a penalty term using the existence of a local error bound. For the case that non negativity constraints are additionally present, we consider an ℓ^1-penalty term. This special case occurs, for instance, in portfolio optimization. For this approach, we show that Karush-Kuhn-Tucker points of a penalised problem, for increasing penalty parameters, fulfil a necessary optimality condition for the reformulation in the limit.
Furthermore, we consider the application of a sequential quadratic programming (SQP) method to the reformulation by using a piecewise decomposition of the feasible set. Our theoretical results yield an explanation for the results delivered by a (standard) SQP solver applied to the reformulation.
Moreover, we consider regularisation methods for the reformulation. We prove convergence of a Scholtes-type regularisation and an exponential regularisation. The former already proved to perform well numerically for MPCCs. Using the second order optimality conditions, we expand the convergence theory for this method. We then show how the approach can be used to expand the convergence theory of a whole class of regularisation methods.
Subsequently we present and discuss computational results. We use the reformulation as a

model for sparse portfolios, which we construct using historical market data. For various time spans these portfolios yield a better Sharpe-ratio than an equal-weight portfolio.
Additionally, we present computational results of the numerical methods for the reformulation. We test the methods for portfolio optimization problems using different risk measures and a range of test cases. The application of the regularisation methods yields better results than a commercial solver for nonlinear optimization problems. Among the penalty methods, the ℓ^1-penalty method yields the best results. If the objective is to compute a good solution in a short amount of time, the Scholtes-type regularisation compares favourably against a commercial global solver, as further numerical results indicate.

Zusammenfassung

In einer Vielzahl von Anwendungen spielen nichtlineare restringierte Optimierungsprobleme eine Rolle. Gegenstand dieser Arbeit sind nichtlineare Optimierungsprobleme mit einer *Kardinalitätsrestriktion*. Durch diese Nebenbedingung wird die Anzahl der Komponenten eines zulässigen Vektors, die ungleich Null sind, beschränkt. Zu den Anwendungen von Kardinalitätsrestriktionen zählen unter anderem Portfolio-Optimierung, Compressed Sensing, sowie logistische Problemstellungen.

Da die Kardinalitätsrestriktion durch eine unstetige Funktion beschrieben wird, ist die Anwendung von herkömmlichen Verfahren aus der nichtlinearen Optimierung nicht ohne weiteres möglich. Eine Möglichkeit ist die Anwendung von Verfahren aus der diskreten Optimierung auf eine Umformulierung mit Binärvariablen.

In dieser Arbeit befassen wir uns mit einem anderen Ansatz von Burdakov et al. (2016): Mittels kontinuierlicher Hilfsvariablen wird die Kardinalitätsrestriktion umformuliert, wodurch auch nichtlineare Probleme abgedeckt werden. Die Umformulierung weist Ähnlichkeit zu einem Optimierungsproblem mit Komplementaritätsnebenbedingungen (MPCC) auf und verletzt ebenfalls Bedingungen (Constraint Qualifications), unter denen Optimalitätsbedingungen erster Ordnung in einem lokalen Minimum gelten. Aus diesem Grund wurden von Burdakov et al. (2016) und Červinka et al. (2016) angepasste Optimalitätsbedingungen erster Ordnung hergeleitet.

Diese Arbeit knüpft an diese Ergebnisse an und enthält neue Optimalitätsbedingungen zweiter Ordnung für die Umformulierung. Diese gelten ebenfalls unter angepassten Constraint Qualifications, von denen man, im Gegensatz zu herkömmlichen Constraint Qualifications, annehmen kann, dass sie für die Umformulierung erfüllt sind. Wir leiten eine notwendige Optimalitätsbedingung zweiter Ordnung, eine hinreichende Optimalitätsbedingung zweiter Ordnung sowie ein Ergebnis zur Eindeutigkeit stationärer Punkte her. Zusätzlich formulieren wir diese Bedingungen bezüglich des ursprünglichen Problems. Diese Optimalitätsbedingungen können beispielsweise benutzt werden, um zu überprüfen, ob ein stationärer Punkt tatsächlich ein lokales Minimum ist. Für die hinreichende Optimalitätsbedingung zweiter Ordnung verwenden wir eine schwächere Voraussetzung als in der Theorie für herkömmliche nichtlineare Optimierungsprobleme. Ferner wird die Existenz einer lokalen Fehlerschranke für die Umformulierung hergeleitet. Sowohl die Optimalitätsbedingungen zweiter Ordnung als auch das Ergebnis zur Existenz einer lokalen Fehlerschranke werden anschließend für die Konvergenztheorie numerischer Verfahren verwendet.

Diese numerischen Verfahren stellen einen weiteren Schwerpunkt der Arbeit dar. Mit Hilfe des Ergebnisses zur Existenz einer lokalen Fehlerschranke leiten wir einen exakten Strafterm her. Für den Fall, dass Nichtnegativitätsnebenbedingungen vorhanden sind, betrachten wir zudem einen ℓ^1-Strafterm. Dieser Spezialfall ist zum Beispiel in der Portfolio-Optimierung von Interesse. Wir zeigen, dass Karush-Kuhn-Tucker Punkte eines Hilfsproblems, welches den ℓ^1-Strafterm verwendet, für wachsende Strafparameter im Grenzübergang eine notwendige Optimalitätsbedingung für die Umformulierung erfüllen.

Zusätzlich untersuchen wir die Anwendung eines Sequential Quadratic Programming (SQP)

Verfahrens auf die Umformulierung theoretisch, wofür wir eine Zerlegung des ursprünglichen Optimierungsproblems verwenden. Unsere Ergebnisse liefern eine mögliche Erklärung für das Konvergenzverhalten eines SQP Verfahrens, welches wir im letzten Kapitel der Arbeit beobachten.
Darüber hinaus übertragen wir Regularisierungsverfahren auf die Umformulierung. Wir zeigen die Konvergenz einer Scholtes-artigen Regularisierung und einer Exponential Regularisierung. Erstere lieferte bereits für MPCCs gute numerische Ergebnisse. Mit Hilfe der Optimalitätsbedingungen zweiter Ordnung bauen wir die Konvergenztheorie für diese Regularisierung weiter aus und zeigen, wie man dieses Vorgehen auf weitere Regularisierungsverfahren übertragen kann.
Anschließend präsentieren wir numerische Ergebnisse. Wir verwenden die Umformulierung als Modell für dünnbesetzte Portfolios, die wir mit Hilfe historischer Kursentwicklungen konstruieren. Diese Portfolios weisen für verschiedene Zeiträume ein höheres Sharpe-Ratio als ein gleichverteiltes Portfolio auf. Die Ergebnisse sprechen für die Umformulierung als Modell für dünnbesetzte Portfolios.
Zusätzlich werden numerische Ergebnisse der Verfahren für die Umformulierung präsentiert und diskutiert. Wir testen die Verfahren an Portfolio-Optimierungsproblemen mit verschiedenen Zielfunktionen und für eine Reihe von Testfällen. Die Anwendung von Regularisierungsverfahren liefert im Großteil der Fälle bessere Ergebnisse als ein kommerzieller Löser für nichtlineare Optimierungsprobleme. Unter den Strafkostenansätzen ist der ℓ^1-Strafterm vielversprechend, welcher sich für bestimmte Zielfunktionen ebenfalls gegen den kommerziellen Löser durchsetzen kann. Gegeben der Zielsetzung eine gute (aber nicht notwendigerweise die globale) Lösung in möglichst kurzer Zeit zu berechnen, vergleichen wir abschließend Ergebnisse der Scholtes-artigen Regularisierung mit Ergebnissen eines Lösers für gemischt-ganzzahlige Probleme.

Contents

1 Introduction

Practical and theoretical questions from a vast range of areas in science and industry can be modelled using a nonlinear constrained optimization problem. This thesis deals with a certain type of optimization problems: *Cardinality constrained optimization problems.*
Let $n, m, p \in \mathbb{N}$, $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^p$. We consider cardinality constrained optimization problems of the form

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{s.t.} \quad g(x) \le 0,\ h(x) = 0, \qquad \|x\|_0 \le \kappa, \tag{1.1}$$

where $\kappa \in \mathbb{N}$ with $0 \le \kappa < n$. The *cardinality constraint* $\|x\|_0 \le \kappa$ is of our special interest. For $x \in \mathbb{R}^n$ the mapping $\|\cdot\|_0$ is defined as

$$\|x\|_0 := |\operatorname{supp}(x)| = |\{i \in \{1,\dots,n\} \,:\, x_i \neq 0\}|.$$

This constraint limits the number of nonzero elements of a feasible vector x of (1.1) to at most κ. The functions f, g and h are possibly nonlinear and throughout this thesis we assume them to be continuously differentiable. Whenever we need them to be twice continuously differentiable, we state this explicitly.
Cardinality constraints have been studied in the context of sparse portfolio optimization where they limit the number of active positions in a portfolio [10]. Further applications are the subset selection problem in regression [63], support vector machines [86], as well as the compressed sensing technique [16]. Cardinality constraints also play a role in a range of logistic or planning problems such as cash management in automatic teller machines [35], lot sizing [34], emergency medical services [67] or network design [83].
Although it is sometimes referred to as *zero norm*, the mapping $\|\cdot\|_0$ is not a norm. In fact it is not even continuous. This makes (1.1) hard to solve: Even if the mappings f, g and h are are continuously differentiable, methods from nonlinear optimization cannot be applied directly. Moreover, testing feasibility for (1.1) is known to be NP-complete [10].
A recent approach is the reformulation of the cardinality constraint using complementarity constraints. We consider the following reformulation of (1.1):

$$\begin{aligned} \min_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n} f(x) \quad \text{s.t.} \quad & g(x) \le 0, \quad h(x) = 0, \\ & 0 \le y \le e, \quad e^T y \ge n - \kappa, \\ & x \circ y = 0. \end{aligned} \tag{1.2}$$

We denote the Hadamard product, i.e. component-wise product, of two vectors $a, b \in \mathbb{R}^n$ by $a \circ b \in \mathbb{R}^n$. The vector $e \in \mathbb{R}^n$ represents the vector whose components are units, i.e. $e = (1,\dots,1)^T \in \mathbb{R}^n$. Throughout this thesis we will also use $e_i \in \mathbb{R}^n$ to denote the i-th unit vector in $\mathbb{R}^n$. Let

$$Z := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \,:\, (x,y) \text{ is feasible for } (1.2)\}.$$

be the feasible set of (1.2). We will refer to (1.2) as *complementarity formulation*. The above nonlinear program is the main focus of this thesis.
The auxiliary variable y can be seen as a counter for the zero elements of the vector x. For every $i = 1, \dots, n$, due to the constraint $x_i \cdot y_i = 0$, a component y_i can only be positive, if $x_i = 0$. Since the constraint $e^T y \geq n - \kappa$ is fulfilled at the same time, we know that x has at least $n - \kappa$ components that are equal to zero. Consequently x fulfils the cardinality constraint. Using an auxiliary variable that counts the *zero* elements of x goes back to [47]. In [14] and [26] the relation between global and local solutions of (1.1) and (1.2) was established. We will discuss these results in the following chapter in detail.
In contrast to the mixed-integer reformulation in [10], problem (1.2) has only continuous variables. This potentially allows the application of methods from nonlinear optimization and therefore covering nonlinear problems. However, due to the complementarity constraint, most conditions that ensure that optimality conditions hold in a local minimum cannot be expected to be satisfied, see [15]. The complementarity formulation has a very similar structure to a mathematical program with complementarity constraints (MPCC). In that setting the additional constraint $x \geq 0$ is present. These problems also violate most standard constraint qualifications. Therefore custom constraint qualifications, stationary conditions and numerical methods for MPCCs were introduced. Yet, even if the additional constraint $x \geq 0$ is present, the results for MPCCs are not readily applicable, since the feasible set of (1.2) violates most of the custom constraint qualifications for MPCCs, see [15]. In [14, 15] concepts from the theory on MPCCs were transferred to the complementarity formulation. Some of the results for the complementarity formulation are even stronger than the corresponding results for MPCCs. In this thesis we further follow this path. For an overview of the subject of MPCCs, see [59, 69] (whenever we resort to particular concepts, we will give further references).
An earlier reformulation of cardinality constraints was introduced in [10] using binary auxiliary variables for the case of polyhedral constraints and a quadratic objective function. This formulation lead to the application of methods from discrete optimization: In [9] a branch & bound method for a quadratic objective function was proposed. In [58] a concave reformulation is considered and applied to portfolio selection problems. Approximation techniques such as simulated annealing are studied in [17, 78]. For a convex objective function, an approximation of the cardinality constraint with the ℓ^1-norm is studied in [89], see also the survey article [84]. Linear programs with (multiple) overlapping cardinality constraints are considered in [27]. Cardinality constraints were first motivated by an application to portfolio optimization in [10] and since then have been further studied in this context. In [66] a local relaxation method was proposed. The application of a convex penalty function was considered in [25].
Compared to the literature covering mixed-integer or approximation methods for cardinality constrained optimization problems, there are few publications covering approaches from nonlinear optimization. In [6] a special case is considered, in which no further constraints, except the cardinality constraint, are present. For this case, optimality conditions from nonlinear optimization and algorithms are investigated. In [71] first and second order optimality conditions are given. These are formulated using the original cardinality constraint and use suitable normal cones of the corresponding feasible set. After being introduced in [14, 26], the complementarity formulation (1.2) was further studied in [15, 88, 11, 13].

The outline of this thesis is as follows. Chapter 2 contains a brief review of two applications of cardinality constrained optimization problems. Among them is a portfolio selection problem which we will consider again in the chapter covering numerical results. Furthermore, in Chapter 2, we discuss the relation between solutions of the cardinality constrained optimization problem (1.1) and solutions of the complementarity formulation (1.2). While these results were originally established in [14], we use a broader setting which includes partial and multiple cardinality constraints.

In Chapter 3 we study theoretical results for the complementarity formulation. Section 3.1 is a brief review of optimality conditions for standard nonlinear programs. In Section 3.2 we discuss why standard constraint qualifications from nonlinear optimization cannot be expected to hold for the complementarity formulation. We then consider custom constraint qualifications and first order optimality conditions to overcome this problem. This was originally done in [14, 15].

In Section 3.3 we present second order optimality conditions for (1.2): We prove both a necessary and a sufficient second order optimality condition for S-stationary points, which complement the first order optimality conditions. For M-stationary points, we prove local uniqueness regarding the variable x of the original problem (1.1) also using a second order condition. Additionally, for all of these three results we provide a formulation in terms of the original problem (1.1) only. The second order optimality conditions expand the set of optimality conditions for (1.2) which hold under custom constraint qualifications. Moreover, these results are central in expanding the convergence theory of regularisation schemes which we study in Section 4.3 and play also a role for a piecewise sequential quadratic programming (SQP) approach in Section 4.2. The results on second order optimality conditions are from [13], together with Alexandra Schwartz.

In Section 3.4 we address two related approaches for optimality conditions for cardinality constrained optimization problems. We consider optimality conditions for the unconstrained case from [6], as well as the first and second order optimality conditions from [71], which are derived for the cardinality constrained optimization problem. We discuss relations between these optimality conditions and the optimality conditions for the complementarity formulation.

In Section 3.5 we use a piecewise decomposition of the complementarity formulation to prove the existence of a local error bound. This result plays again a role in Section 4.1, where we use it to prove exactness of a distance-based penalty function. The local error bound result is from [12], together with Christian Kanzow and Alexandra Schwartz.

In Chapter 4 we study three classes of numerical methods for the complementarity formulation. Section 4.1 contains two penalty approaches. The first approach uses a distance-based penalty function which is exact provided a custom constraint qualification holds. The second approach is applicable to the special case of (1.1) in which we have the additional constraint $x \geq 0$. This case is of interest for example in portfolio optimization. We use an ℓ^1-norm penalty term to move the complementarity constraint of (1.2) to the objective function and show that the limit of KKT points of this partially penalised problem, provided it is feasible, fulfils custom first order optimality conditions for (1.2). The results on penalisation techniques are from [12], together with Christian Kanzow and Alexandra Schwartz.

In Section 4.2 we consider a piecewise SQP scheme for the cardinality constrained optimization problem. An application of an SQP method to the complementarity formulation yields quadratic subproblems which correspond to quadratic subproblems of a certain decomposition of the feasible set of (1.1). Using this decomposition, we then investigate the behaviour of a (standard) SQP method applied to the complementarity formulation.

In Section 4.3 we consider regularisation methods for the complementarity formulation. This

class of methods has been studied in the context of MPCCs. By relaxing the complementarity constraint of (1.2) one obtains a regularised nonlinear program which can be expected to fulfil standard constraint qualifications. The regularisation methods then compute KKT points of a sequence of such regularised programs, and aim to obtain a point which fulfils a first order optimality condition in the limit.
We consider three regularisation methods: Firstly, we consider a Scholtes-type regularisation. This type of regularisation was among the first to be studied for MPCCs [80]. For the complementarity formulation, we obtain a convergence result which is stronger than the corresponding result for MPCCs. Furthermore, we prove the existence of local solutions of the regularised programs. Moreover, using the second order conditions from Chapter 3, we are able to prove a uniqueness result for the limit points. The results on the Scholtes-type regularisation are from [11, 13], together with Martin Branda, Michal Červinka and Alexandra Schwartz.
Secondly, we consider the Kanzow-Schwartz regularisation from [54], which was successfully adapted to the complementarity formulation in [14]. We expand the convergence results for this method analogously to the Scholtes-type regularisation using the second order optimality conditions from Chapter 3.
Thirdly, we consider a regularisation that uses an exponential function to relax the complementarity constraint. This scheme was studied in the context of chance constrained optimization problems in [1]. For this method we are able to prove a convergence result analogous to the Scholtes-type regularisation.
In Chapter 5 we examine numerical results for the complementarity formulation. In the first part, Section 5.1, we use (1.2) as a model for sparse portfolio selection. We use historical stock market data for the model and use the Scholtes-type regularisation to solve the resulting complementarity formulation. The constructed portfolios can compete with an evenly distributed portfolio, which is considered a tough benchmark, in terms of the Sharpe ratio. These results indicate that the complementarity formulation serves well as a model for sparse portfolios.
In Section 5.2 we test the penalty and regularisation methods from Chapter 4 on sparse portfolio optimization problems. We apply the methods to a range of test problems using different risk measures such as value-at-risk or conditional value-at-risk. We compare these methods with a solver for nonlinear optimization problems. Furthermore, we compute solutions of a mixed-integer formulation of the portfolio optimization problem with Gurobi. We use the obtained results as a reference to evaluate the Scholtes-type regularisation's performance.

Before we proceed, let us state a few more standard symbols that we will use throughout this thesis: We use $\mathbb{R}_+ = [0, \infty)$ for the set of nonnegative real numbers. We denote the complement of a set A by A^C. For $A \subseteq \mathbb{R}^n$ we denote the orthogonal projection onto the set A by P_A and its convex hull by $\operatorname{conv}(A)$. For an arbitrary norm on $\mathbb{R}^n$, we denote by $B_r(x)$ the open ball of radius r with centre at the point x, its closure by $\overline{B_r(x)}$ and its boundary by $\partial B_r(x)$. As mentioned before, the unit vectors are denoted by $e_i \in \mathbb{R}^n$ and $e \in \mathbb{R}^n$ is the vector consisting of units. For two vectors $a, b \in \mathbb{R}^n$ we denote the Hadamard product, i.e. the component-wise product, by $a \circ b \in \mathbb{R}^n$ and we denote the line segment connecting the points a and b by $[a, b]$.

2 Cardinality Constrained Optimization Problems: Applications and a Reformulation

In this chapter we first give a brief overview of two applications of cardinality constrained optimization problems in Section 2.1. We consider a portfolio optimization problem for sparse portfolio selection. Additionally we consider a model from communications engineering and expand it by introducing cardinality constraints.
In Section 2.2 we study the relation between solutions of the cardinality constrained optimization problem and the complementarity formulation. We consider a slightly broader setting which includes partial and multiple (possibly overlapping) cardinality constraints.

2.1 Applications of Cardinality Constrained Optimization Problems

Cardinality constraints play a role in a range of applications such as the subset selection problem in regression [63], support vector machines [86], compressed sensing [16], cash management in automatic teller machines [35], lot sizing [34], emergency medical services [67] or network design [83]. Generally, cardinality constraints can be used to count the nonzero elements of a continuous variable without introducing a binary variable for this purpose. In sparse portfolio optimization the continuous variable models an investment in certain assets while the cardinality constraint counts the number of assets in the portfolio. Likewise, in the antenna placement problem which we discuss later, the continuous variable can represent a maximum data rate, while the cardinality constraint limits the number of connections provided by an antenna.

Portfolio Optimization

Portfolio optimization is the application that motivated the study of cardinality constraints in [10]. We consider the following portfolio optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \ f(x) \quad \text{s.t.} \quad & e^T x = 1, \\ & l \leq x \leq u, \\ & \|x\|_0 \leq \kappa. \end{aligned} \tag{2.1}$$

In this scenario an investor can choose from n financial assets. A component x_i represents the investment in asset $i \in \{1, \dots, n\}$. The objective is to minimise a risk measure $f : \mathbb{R}^n \to \mathbb{R}$ while distributing the whole budget, which is modelled by the first constraint. The investment is bounded by lower and upper bounds $l, u \in \mathbb{R}^n$. The cardinality constraint limits the number of non-zero components of x to $\kappa < n$. This means that there can be at most κ active positions in a portfolio. Thus portfolios obtained from this model are easier to manage and transaction costs are reduced. The assets' returns are assumed to be randomly distributed. A possible risk measure, used in the classic model by Markowitz [61], is the variance $f(x) = x^T Qx$,

where Q is the covariance matrix of the returns. In practice, the variance is estimated from historical stock market data, which exposes the model to estimation errors. Limiting the number of active positions in the portfolio can have a stabilising effect and therefore result in better performing portfolios, see [37].
We will consider portfolio optimization problems in Chapter 5 where we evaluate the performance of sparse portfolios obtained by using (1.2) as a model. Furthermore, we will test the performance of different numerical methods for (1.2) on a set of test problems and using different types of risk measures such as value-at-risk or conditional value-at-risk.

Extension of an Antenna Placement Problem

Next, we consider a task from communications engineering. For the design of a wireless network, its layout and the distribution of bandwidth among wireless devices are to be determined. In [70] a model for the placement and orientation of antennas on base stations in a wireless network is motivated and discussed. A number of wireless devices should be connected to antennas which are placed on base stations. The objective is to minimise the total number of antennas in use while connecting all devices to an antenna. The model considers a static case in which the wireless devices are not moving. This case can be found for example in universities or hospitals.
In the original model there are two types of *integer* variables present: A vector x models the decision to serve a given wireless device while a vector y models the placement of the antennas at possible locations. In our modification of the model we use a *continuous* variable x which models the data rate provided by an antenna to a given wireless device. If the variable, e.g. the provided data rate, is zero, there is no connection. We model this with cardinality constraints. To formally state the model we introduce the parameters in Table 2.1, cf. [70].
To be connected to the network a device j must be within reach of a base station i. In this model the base stations serve as gateways to the network. The antennas placed at these base stations are orientated such that they cover a certain spatial sector of the area. The set $I(j)$ contains all base stations within reach of device j and the label l_{ij} is used to for the sector of base station i where device j is located.
Let the vector $x^k = (x_1^k, \dots, x_n^k)$, $k \in K$, contain the bandwidth assigned to each device by antenna k. We combine these in the vector

$$x = (x^1, \dots, x^m) \in \mathbb{R}^{m \cdot n}.$$

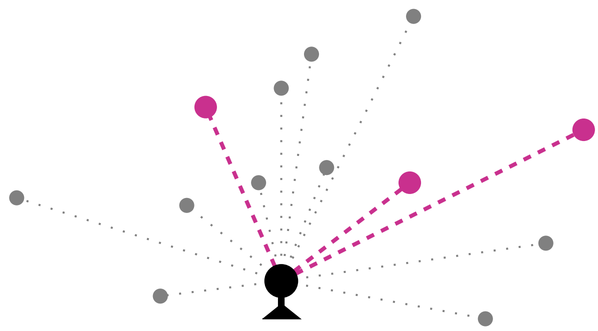


Figure 2.1: Illustration of an antenna providing access for wireless devices to the network.

Index sets

$I = \{1, \ldots, o\}$	set of base stations
$L = \{1, \ldots, p\}$	set of sectors
$J = \{1, \ldots, n\}$	set of wireless devices
$K = \{1, \ldots, m\}$	set of antennas

Static parameters

$b(j)$	bandwidth of wireless device j
$\beta(k)$	total bandwidth that can be served by antenna k, we assume $\beta(k) > b(j)\ \forall k, \forall j$
θ	angle span of antenna in the unit sector
c	maximum number of devices per antenna

$$m_{jil} = \begin{cases} 1, & \text{device } j \text{ is located at sector } l \text{ base station } i, \\ 0, & \text{otherwise.} \end{cases}$$

Variable parameters:

$x_j^k \in [0, b(j)]$ bandwidth of connection by antenna k to device j

$$y_{il}^k = \begin{cases} 1, & \text{if antenna } k \text{ is located on base station } i \\ & \quad \text{and covering sectors } l \text{ to } (l+\theta) \mod |L| \\ 0, & \text{otherwise,} \end{cases}$$

Table 2.1: An Overview over the variables and constants in the antenna placement problem.

The objective of the original antenna location problem is to minimise the total number of antennas serving the wireless devices. Using the variable x to additionally model the bandwidth of a connection, we use the objective to also maximise the total bandwidth in the network. The weight between the total bandwidth and the number of antennas used can be adjusted by the parameter $w \geq 0$. The antenna location problem is then given by

$$\begin{aligned} \min_{x,y} \quad & \sum_{k,i,l} y_{il}^k - w \cdot \sum_{k,j} x_j^k \\ \text{s.t.} \quad & y_{il}^k \in \{0,1\}, && \forall k \in K, \forall i \in I, \forall l \in L, \\ & x_j^k \in [0, b(j)], && \forall j \in J, \forall k \in K, \\ & x_j^k \leq b(j) \cdot \sum_{i \in I(j)} \sum_{l_{ji}-\theta \leq l \leq l_{ji}} y_{il}^k, && \forall j \in J, \forall k \in K, && (2.2) \\ & \sum_j x_j^k \leq \beta(k), && \forall k \in K, && (2.3) \\ & \sum_{i,l} y_{il}^k \leq 1, && \forall k \in K, && (2.4) \\ & \|(x_1^k, \ldots, x_n^k)\|_0 \leq c, && \forall k \in K, && (2.5) \\ & \|(x_j^1, \ldots, x_j^m)\|_0 = 1, && \forall j \in J. && (2.6) \end{aligned}$$

For a connection, i.e. $x_j^k > 0$, a device j must be reachable by antenna k, this is modelled by constraint (2.2). The constraint on the bandwidth an antenna k can serve is given by

(2.3). Constraint (2.4) models that each antenna can be placed at most once (to a certain base station i oriented to a certain sector l). We have two sets of cardinality constraints in the model. Firstly, (2.5) models the maximum number of devices assigned to one antenna. Secondly, (2.6) models the fact that each device should be assigned to exactly one antenna. We call these constraints *partial* cardinality constraints, as they restrict the number of nonzero components of a subset of the components of x.
The antenna location problem in this form is a mixed-integer program. Relaxation of the binary variables y_{il}^k, for example for a branch and bound method, would result in a nonlinear program with cardinality constraints. Another possible application of the model is the minimisation of a social welfare function to obtain a desired bandwidth allocation.

2.2 A Continuous Reformulation

In this section we study the relation between the cardinality constrained optimization problem (1.1) and the complementarity formulation (1.2), which has been established in [14] and [26]. In the latter reference it is the so called half-complementarity formulation.
For the discussion of the relation between global and local minima we consider a broader setting. As before, let $f\colon \mathbb{R}^n \to \mathbb{R}$, $g\colon \mathbb{R}^n \to \mathbb{R}^m$, $h\colon \mathbb{R}^n \to \mathbb{R}$. Throughout this thesis we will use the following sets:

$$\begin{aligned} X &:= \{x \in \mathbb{R}^n \,:\, g_i(x) \le 0,\ i = 1, \dots, m,\ h_i(x) = 0,\ i = 1, \dots, p\}. \\ \mathcal{X} &:= \{x \in X \,:\, \|x\|_0 \le \kappa\}. \end{aligned}$$

To model partial or multiple cardinality constraints, let $S_1, \dots, S_q$ be arbitrary subsets of $\{1, \dots, n\}$ and $0 < \kappa_j < |S_j|$ for $j = 1, \dots, q$. We consider the optimization problem

$$\begin{aligned} \min_x\ f(x) \quad \text{s.t.} \quad & x \in X, \\ & \|x_{S_j}\|_0 \le \kappa_j, \quad j = 1, \dots, q, \end{aligned} \tag{2.7}$$

where the vector x_{S_j} is given by $x_{S_j} = (x_j)_{i \in S_j}$ for $j = 1, \dots, q$. Overlapping cardinality constraints are included in this setting, i.e. we allow $S_j \cap S_k \neq \emptyset$ for $j \neq k$. Hence we will refer to (2.7) as *optimization problem with overlapping cardinality constraints*. Analogous to the results in [14, Section 3] problem (2.7) can be reformulated in to a continuous optimization problem. We consider two reformulations of (2.7). The first is a mixed-integer formulation given by

$$\begin{aligned} \min_{x,y}\ f(x) \quad \text{s.t.} \quad & x \in X, \\ & \sum_{i \in S_j} y_i \ge |S_j| - \kappa_j, \qquad j = 1, \dots, q, \\ & y_i \in \{0, 1\}, \qquad\qquad\quad i = 1, \dots, n, \\ & x_i \cdot y_i = 0, \qquad\qquad\quad\ \ i = 1, \dots, n. \end{aligned} \tag{2.8}$$

The auxiliary variable y can be seen as a counter for the zero components of x: Because of the constraint $x_i \cdot y_i = 0$, the component y_i can only be 1 if x_i is zero. On the other hand, for each $j \in \{1, \dots, q\}$ the constraint $\sum_{i \in S_j} y_i \ge |S_j| - \kappa_j$ has to be fulfilled. Therefore we know that at least $|S_j| - \kappa_j$ elements of x_{S_j} have to be zero. Thus for a vector (x, y) feasible for

(2.8), the x part fulfils each of the cardinality constraints of (2.7). This argumentation will be formalised to prove the results to follow.
The second reformulation we are considering is the continuous relaxation of (2.8), given by

$$\begin{aligned}\min_{x,y} f(x) \quad \text{s.t.} \quad & x \in X, \\ & \sum_{i \in S_j} y_i \geq |S_j| - \kappa_j, && j = 1, \dots, q, \\ & 0 \leq y_i \leq 1, && i = 1, \dots, n, \\ & x_i \cdot y_i = 0, && i = 1, \dots, n.\end{aligned} \tag{2.9}$$

To obtain (2.9), the binary variable y is relaxed to a continuous variable $y \in \mathbb{R}^n$. The result is an optimization problem with only continuous variables. Figure 2.2 illustrates the complementarity formulation with integer and continuous variables.

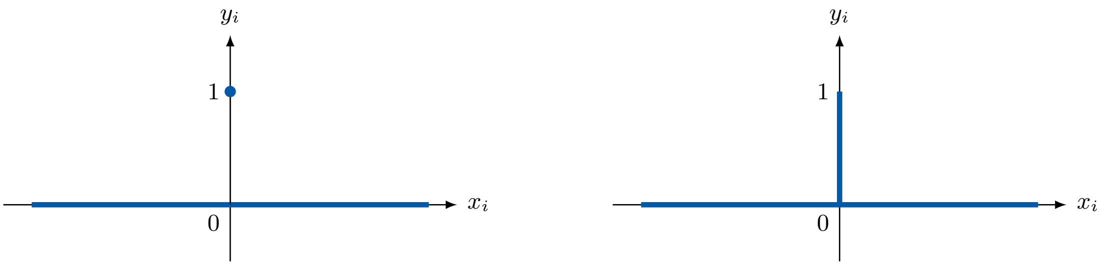


Figure 2.2: Pairs (x_i, y_i) (in blue) that fulfil the the complementarity constraint of the mixed-integer formulation (2.8) and the relaxed complementarity constraint of (2.9).

We begin with the relation between global minima of (2.7) and global minima of the mixed-integer formulation (2.8). The lines of argument for the following results where established in [14].

Theorem 2.1. *A vector $x^* \in \mathbb{R}^n$ is a solution of* (2.7) *if and only if there exists a vector y^* such that (x^*, y^*) is a solution of the reformulation* (2.8).

Proof. Since both optimization problems, (2.7) and (2.8), have the same objective function, we only have to show feasibility. Let x be feasible for (2.7) and let

$$y_i := \begin{cases} 0, & \text{if } x_i \neq 0, \\ 1, & \text{if } x_i = 0. \end{cases}$$

Then by construction of y we have $y_i \in \{0,1\}$ and $x_i \cdot y_i = 0$ for all $i = 1, \dots, n$. Since $\|x_{S_j}\|_0 \leq \kappa$ for all $j = 1, \dots, q$, we also have $\sum_{i \in S_j} y_i \geq |S_j| - \kappa_j$ for all $j = 1, \dots, q$. Thus (x, y) is feasible for (2.8). Now let (x, y) be a feasible point of (2.8). Then $y_i \in \{0,1\}$, $x_i \cdot y_i = 0$ for $i = 1, \dots, n$ and $\sum_{i \in S_j} y_i \geq |S_j| - \kappa_j$. For $j = 1, \dots, q$ define

$$I_j := \{i \in S_j \, : \, y_i = 1\}$$

We have $y_i \in \{0,1\}$ and thus $|I_j| = \sum_{i \in S_j} y_i \geq |S_j| - \kappa_j$ for all $j = 1, \dots, q$. Because $x_i = 0$ for all $i \in I_j \subseteq S_j$, we have $\|x_{S_j}\|_0 \leq |S_j| - |I_j| \leq |S_j| - (|S_j| - \kappa_j) = \kappa_j$, consequently x is feasible for (2.7). $\square$

Next, we consider the relations between global and local minima of (2.7) and the relaxation (2.9).
The cardinality constrained problem (1.1) is a special case of (2.7), with $q = 1$ and $S_1 = \{1, \ldots, n\}$. In this case, the complementarity formulation (1.2) coincides with (2.9). Thus the following two results hold for (1.1) and the complementarity formulation (1.2) as well.

Theorem 2.2. *A vector $x^* \in \mathbb{R}^n$ is a solution to* (2.7) *if and only if there exists a vector y^* such that (x^*, y^*) is a solution of the relaxed problem* (2.9)*.*

Proof. Since both optimization problems, (2.7) and (2.9), have the same objective function, we only have to show feasibility.

"$\Rightarrow$": Let $x^* \in X$ be a global solution of (2.7). Then we have $\|x^*_{S_j}\|_0 \leq \kappa_j$ for all $j = 1, \ldots, q$. Let

$$y_i^* := \begin{cases} 1, & \text{if } x_i^* = 0, \\ 0, & \text{if } x_i^* \neq 0. \end{cases}$$

Then we have $y^* \in [0, 1]$ and $y_i^* \cdot x_i^* = 0$ for all $i = 1, \ldots, n$ and

$$\sum_{i \in S_j} y_i^* = |S_j| - \|x^*_{S_j}\|_0 \geq |S_j| - \kappa_j$$

for all $j = 1, \ldots, q$. Hence (x^*, y^*) is feasible for (2.9).

"$\Leftarrow$": Let (x^*, y^*) be a global solution of (2.9). Then we have $x^* \in X$ and $y_i^* \in [0, 1]$ and $x_i^* \cdot y_i^* = 0$ for $i = 1, \ldots, n$ as well as $\sum_{i \in S_j} y_i^* \geq |S_j| - \kappa_j$ for all $j = 1, \ldots, q$. Let

$$I_j := \{i \in S_j \,:\, y_i \in (0, 1]\}, \quad j = 1, \ldots, q.$$

Because of $y_i^* \in [0, 1]$ we have $|I_j| \geq \sum_{i \in S_j} y_i^* \geq |S_j| - \kappa_j$, hence

$$\kappa_j \geq |S_j| - |I_j|$$

for all $j = 1, \ldots, q$. We have $x_i^* = 0$ for all $i \in I_j$ and all $j = 1, \ldots, q$ and consequently

$$\|x^*_{S_j}\|_0 \leq |S_j| - |I_j| \leq \kappa_j$$

for all $j = 1, \ldots, q$. Hence x^* is feasible for (2.7).

□

The previous result gives a one to one relation between global minima of the optimization problem with overlapping cardinality constraints and its mixed-integer reformulation. For the relation between local minima the following result holds.

Theorem 2.3. *Let $x^* \in \mathbb{R}^n$ be a local minimum of problem* (2.7)*. Then there exists a y^* such that (x^*, y^*) is a local minimum of the relaxed problem* (2.9)*.*

Proof. Let x^* be a local solution of (2.7). Then we have $x^* \in X$, $\|x^*_{S_j}\|_0 \leq \kappa_j$ for all $j = 1, \ldots, q$ and there exists a radius $r_1 > 0$ such that

$$f(x) \geq f(x^*)$$

for all $x \in B_r(x^*) \cap \{x \in X : \|x_{S_j}\|_0 \leq \kappa_j \ \forall j = 1, \ldots, q\}$. Let

$$y_i^* := \begin{cases} 1, & \text{if } x_i^* = 0, \\ 0, & \text{if } x_i^* \neq 0. \end{cases}$$

Then we have $0 \leq y_i^* \leq 1$ and $x_i^* \cdot y_i^* = 0$ for all $i = 1, \ldots, n$ as well as

$$\sum\nolimits_{i \in S_j} y_i^* \geq |S_j| - \|x_{S_j}^*\|_0 \geq |S_j| - \kappa_j \quad \forall j = 1, \ldots, q.$$

Hence (x^*, y^*) is feasible for (2.9). For all $y \in B_{\frac{1}{2}}(y^*)$ we have $y_i^* = 1 \Rightarrow y_i > 0$ and thus $\{i : y_i = 0\} \subseteq \{i : y_i^* = 0\}$. Let $r_2 := \min\{r_1, \frac{1}{2}\}$ and

$$\begin{aligned}(x, y) \in \mathcal{U} := \big(B_r(x^*) \times B_r(y^*)\big) \cap \big\{(x, y) : \, & x \in X, y_i \in [0, 1], \ x_i \cdot y_i = 0, \forall i = 1, \ldots, n, \\ & \sum_{i \in S_j} y_i \geq |S_j| - \kappa_j, \forall j = 1, \ldots, q\big\}.\end{aligned}$$

be arbitrary. From the definition of y^* we have

$$x_i \neq 0 \Rightarrow y_i = 0 \Rightarrow y_i^* = 0 \Rightarrow x_i^* \neq 0$$

for all $i = 1, \ldots, n$. Hence $x_i^* = 0 \Rightarrow x_i = 0$ and consequently

$$\|x_{S_j}\|_0 \leq \|x_{S_j}^*\|_0 \leq \kappa_j$$

for all $j = 1, \ldots, q$. Thus x is feasible for (2.7) and $x \in B_r(x^*)$, thus $f(x) \geq f(x^*)$ for all $(x, y) \in \mathcal{U}$. □

The converse of the above result is not true in general: For a local minimum (x^*, y^*) of (2.9) the part x^* is not a local minimum of (2.7) in general. A counter example was given in [14, Example 2].

The following result can be seen as a counterpart to Theorem 2.3, in case the cardinality constraint is active. It covers the case of partial cardinality constraints, which can be proven in the same way as [14, Theorem 3.6].

Theorem 2.4 ([14, Theorem 3.6])**.** *Let $q = 1$, let (x^*, y^*) be a local minimum of* (2.9) *and let $\|x_{S_1}^*\|_0 = \kappa_1$ hold. Then x^* is a local minimum of* (2.7)*.*

We now come back to the case in which *one* cardinality constraint is present, for which we can state the following additional relation.

Proposition 2.5 ([14, Proposition 3.5])**.** *Let $(x^*, y^*) \in Z$ be a local minimum of* (1.2)*. Then $\|x^*\|_0 = \kappa$ holds if and only if y^* is unique, i.e. if there is exactly one y^* such that $(x^*, y^*) \in Z$. In this case the components of y^* are binary.*

In the following chapters we will discuss theoretical results, such as optimality conditions, as well as numerical methods for the complementarity formulation (1.2) in-depth. Before we proceed, let us have a look at another possible reformulation of the cardinality constraint.

For this we assume that lower and upper bounds $l, u \in \mathbb{R}^n$ on x are given. By introducing binary auxiliary variables $z \in \{0,1\}^n$, we can reformulate the constraints

$$l \leq x, \leq u, \qquad \|x\|_0 \leq \kappa,$$

to

$$l \circ z \leq x \leq u \circ z, \qquad e^T z \leq \kappa. \tag{2.10}$$

Unlike the auxiliary variable y in (1.2), the variable z can be seen as a counter for the *nonzero* components of x. For pairs (x, z) that fulfil (2.10), we have $z_i = 1$ for each $x_i \neq 0$ and thus $\|x\|_0 \leq e^T z$. This approach goes back to Bienstock [10], see also [33]. For the above reformulation only linear constraints are added, which is an advantage compared to the reformulation (1.2). On the other hand, lower and upper bounds for x need to be known. Moreover, it is easy to see that the relaxation of the binary variable is problematic: For $z \in [0,1]^n$ the estimate $\|x\|_0 \leq e^T z$ does not necessarily hold anymore. See also Figure 2.3 for an illustration. Therefore, for this approach, one has to solve a mixed-integer program.

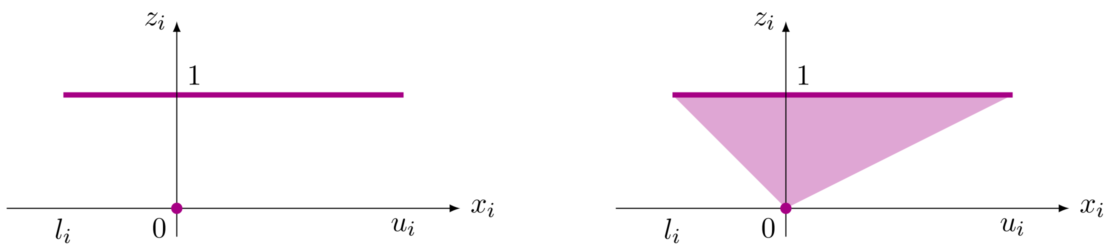


Figure 2.3: Pairs (x_i, z_i) (in purple) that fulfil the reformulation of the cardinality constraint according to (2.10) (left diagram) and its relaxation (right diagram).

3 Theoretical Results

In this chapter we present theoretical results for the complementarity formulation of a cardinality constrained optimization problem. Since we focus on problem-tailored optimality conditions, we start in Section 3.1 with a brief recapitulation of classic results on optimality conditions for nonlinear optimization problems. In Section 3.2 we discuss why these classic result cannot be expected to hold for (1.2) and present two problem-tailored first order optimality conditions, originally introduced by Burdakov, Červinka, Kanzow and Schwartz [14, 15].
These are essential for our subsequent discussion of second order optimality conditions in Section 3.3, as well as for the convergence theory of numerical methods in Chapter 4. We derive both a necessary and a sufficient second order optimality condition which complement the first order optimality conditions. The necessary second order optimality condition holds under a problem-tailored constraint qualification. The sufficient second order optimality condition holds for points that fulfil the stronger of the aforementioned stationary conditions, S-stationary points, and captures the lack of curvature of the objective function with respect to the auxiliary variable y. For points fulfilling the weaker stationary condition, M-stationary points, we prove their uniqueness regarding the x variable under a second order condition.
In Section 3.4 we compare these optimality conditions with two recent other approaches for optimality conditions for the cardinality constrained optimization problem.
We conclude this chapter with a result on the existence of a local error bound for the feasible set of (1.2) in Section 3.5. This result is used in Chapter 4 to prove the exactness of a penalty function.

3.1 Background from Nonlinear Programming

In this section we review classic results on optimality conditions for nonlinear optimization problems of the form

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad g(x) \leq 0, \ h(x) = 0. \tag{3.1}$$

Because this standard form of a nonlinear optimization problem is equivalent to (1.1) without the cardinality constraint, we will use the same notation. As before, let the feasible set of (3.1) be

$$X = \{x \in \mathbb{R}^n \, : \, g_i(x) \leq 0, \ i = 1, \dots, m, \ h_i(x) = 0, i = 1, \dots, p\}$$

and $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^p$. Furthermore, let f, g and h be continuously differentiable and X be nonempty. We will also use the index set

$$I_g(x) = \{i \in \{1, \dots, m\} \, : \, g_i(x) = 0\}$$

for active inequality constraints in a feasible point $x \in X$. We consider first and second order optimality conditions for (3.1) and the constraint qualifications required for those conditions to hold. Furthermore, we repeat second order optimality conditions for (3.1). The literature on the classic results presented in this section is extensive. We refer the reader to [36, 68, 85]. In Section 3.2 we will reason that the presented results cannot be expected to hold for (1.2). Yet we can use them for auxiliary nonlinear problems used in the analysis of (1.2).

3.1.1 First Order Optimality Conditions and Constraint Qualifications

We begin with the definition of the *tangent cone*, which can be used as a local approximation of the feasible set of (3.1).

Definition 3.1. Let $A \subseteq \mathbb{R}^n$ be nonempty and $x \in A$. The set

$$\mathcal{T}_A(x) := \big\{ d \in \mathbb{R}^n \, : \exists (x^k)_{k\in\mathbb{N}} \subseteq A, \exists (t^k)_{k\in\mathbb{N}} \subseteq (0,\infty) : \\ x^k \to x \ (k\to\infty),\ t^k \to 0 \ (k \to \infty) \quad \text{and} \quad \frac{x^k - x}{t^k} \to d \ (k\to\infty) \big\}$$

is called *(Bouligand) tangent cone of A at x*. A vector $d \in \mathcal{T}_A(x)$ is called *tangent (vector) to A at x*.

It can be verified easily that the tangent cone is a cone, thus the name is justified.[1] Using the tangent cone, we obtain a first general characterisation of local minima.

Lemma 3.2. *Let $x^* \in X$ be a local minimum of* (3.1). *Then*

$$\nabla f(x^*)^T d \geq 0 \quad \forall d \in \mathcal{T}_X(x^*).$$

The above result states that there is no descent direction in the tangent cone at x^* if x^* is a local minimum. Because of its definition, the tangent cone can be difficult to compute. A more practical approximation of the feasible set X can be formulated using the first derivatives of the functions g and h.

Definition 3.3. Let x be a feasible point of (3.1). The set

$$\mathcal{L}_X^{NLP}(x) := \{ d \in \mathbb{R}^n \, : \, \nabla g_i(x)^T d \leq 0 \ \forall i \in I_g(x), \nabla h_i(x)^T d = 0, i = 1, \dots, p \} \tag{3.2}$$

is called *linearisation cone of X at x*.

It is easy to verify that the linearisation cone is a cone. We further have

$$\mathcal{T}_X(x) \subseteq \mathcal{L}_X^{NLP}(x)$$

for all $x \in X$. The derivation of first order optimality conditions for (3.1) relies on the presumption that $\mathcal{T}_X(x) = \mathcal{L}_X^{NLP}(x)$ holds true, or at least $\mathcal{T}_X(x)^\circ = \mathcal{L}_X^{NLP}(x)^\circ$.[2] Conditions that ensure this equality are called *constraint qualifications*. To define the linearisation cone the gradients of the functions g and h are used, hence most constraint qualifications are requirements on the gradients of these functions. We define a number of common constraint qualifications for nonlinear optimization problems.

[1] A set $A \subseteq \mathbb{R}^n$ is called *cone* if for all $a \in A$ and all $t \geq 0$: $t \cdot a \in A$.

[2] Given a set $A \subseteq \mathbb{R}^n$, the set $A^\circ := \{ v \in \mathbb{R}^n \, : \, a^T v \leq 0 \ \forall a \in A \}$ is called *polar cone of A*.

Definition 3.4. Let x be a feasible point of (3.1). We say that

(a) the *Linear Independence Constraint Qualification (LICQ)* holds at x, if the gradients

$$\nabla g_i(x), i \in I_g(x),\ \nabla h_i(x), i = 1, \dots, p,$$

are linearly independent.

(b) the *Mangasarian-Fromovitz Constraint Qualification (MFCQ)* holds at x, if the gradients $\nabla h_i(x)$, $i = 1, \dots, p$, are linearly independent and there exists a vector $d \in \mathbb{R}^n$, such that

$$\nabla g_i(x)^T d < 0, i \in I_g(x),\ \nabla h(x)^T d = 0.$$

(c) the *Constant Rank Constraint Qualification (CRCQ)* holds at x, if for all subsets $I_1 \subseteq I_g(x)$ and $I_2 \subseteq \{1, \dots, p\}$ such that the gradients

$$\nabla g_i(x), i \in I_1,\ \nabla h_i(x), i \in I_2$$

are linearly dependent, there exists a neighbourhood $N(x)$ of x such that for all $y \in N(x)$ the gradients

$$\nabla g_i(y), i \in I_1,\ \nabla h_i(y), i \in I_2$$

are also linearly dependent.

(d) the *Abadie Constraint Qualification (ACQ)* holds at x, if $\mathcal{T}_X(x) = \mathcal{L}_X^{NLP}(x)$.

(e) the *Guignard Constraint Qualification (GCQ)* holds at x, if $\mathcal{T}_X(x)^\circ = \mathcal{L}_X^{NLP}(x)^\circ$.

Constraint qualifications such as the above play an essential role in the analysis of (3.1). In a local minimum first order optimality conditions hold, provided a constraint qualification is satisfied. Moreover, constraint qualifications are requirements for many convergence results for numerical methods. We will repeat a useful characterisation of MFCQ and define one additional constraint qualification. To do this, we need the concept of positive linear dependence.

Definition 3.5. Let $\{a_1, \dots, a_m\}$ and $\{b_1, \dots, b_p\}$ be two finite set of vectors. We call $\{a_1, \dots, a_m\}$ and $\{b_1, \dots, b_p\}$ *positively linearly dependent*, if there exist multipliers $\alpha \in \mathbb{R}^m$ and $\beta \in \mathbb{R}^p$, such that $(\alpha, \beta) \neq 0$, $\alpha \geq 0$ and

$$\sum_{i=1}^{m} \alpha_i a_i + \sum_{i=1}^{p} \beta_i b_i = 0. \tag{3.3}$$

Otherwise $a_1, \dots, a_m$ and $b_1, \dots, b_p$ are called *positively linearly independent.*

The following characterisation is useful for arguments using the MFCQ. It can be derived using Motzkin's theorem of the alternative (see [60]) inserting the Jacobians of g and h as the appropriate matrices.

Lemma 3.6. *Let $x \in X$. Then x satisfies MFCQ if and only if the gradients*

$$\{\nabla g_i(x), i \in I_g(x)\} \quad \text{and} \quad \{\nabla h_i(x), i = 1, \dots, p\}$$

are positively linearly independent.

With the concept of positive linear dependence the following constraint qualification can be defined.

Definition 3.7. Let x be a feasible point of (3.1). We say that x satisfies the *Constant Positive Linear Dependence Constraint Qualification (CPLD)*, if for all subsets $I_1 \subseteq I_g(x)$ and $I_2 \subseteq \{1, \dots, p\}$, such that the gradients

$$\{\nabla g_i(x), i \in I_1\} \quad \text{and} \quad \{\nabla h_i(x), i \in I_2\}$$

are positively linearly dependent, there exists a neighbourhood $N(x)$ of x such that for all $y \in N(x)$ the gradients

$$\nabla g_i(y), i \in I_1, \ \nabla h_i(y), i \in I_2$$

are linearly dependent.

The constraint qualifications LICQ, MFCQ, ACQ and GCQ can be found in a number of textbooks on nonlinear programming, see for example [36, 68, 85]. The CRCQ was introduced in [50], CPLD was introduced in [75] and proven to be a constraint qualification in [3]. We now state the Karush-Kuhn-Tucker conditions. As we will see in the theorem afterwards, these conditions are first order necessary optimality conditions for a local minimum of (3.1) provided a constraint qualification holds.

Definition 3.8. Let x be feasible for (3.1). We say that the *Karush-Kuhn-Tucker conditions (KKT conditions)* hold in x if there exist coefficients $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p$ such that

$$\begin{aligned} \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) + \sum_{i=1}^{p} \mu_i \nabla h_i(x) = 0, \\ \lambda_i \geq 0, \quad \lambda_i \cdot g_i(x) = 0 \quad \forall i = 1, \dots, m. \end{aligned}$$

In this case x is called *Karush-Kuhn-Tucker point (KKT point)*. The coefficients λ_i, $i = 1, \dots, m$, and μ_i, $i = 1, \dots, p$, as well as the pair (λ, μ) itself are called *Lagrange multipliers* or *KKT multipliers*.

It is more common to call the triple (x, λ, μ) itself KKT point and the point x *stationary*. However, we use the terminology of Definition 3.8 to better distinguish between KKT points and further stationary concepts, which we will encounter in the subsequent study. We now state a central result: Under a constraint qualification a local minimum of (3.1) is a KKT point.

Theorem 3.9. *Let x^* be a local minimum of* (3.1) *and GCQ or a stronger constraint qualification hold at x^*. Then x^* is a KKT point.*

Furthermore, it is easy to verify that under LICQ the KKT multipliers are unique. Considering the characterisation of MFCQ given by Lemma 3.6, the relations between LICQ, MFCQ, CRCQ and CPLD directly follow from their definitions: LICQ implies MFCQ and CRCQ. Both MFCQ and CRCQ imply CPLD. The relation between CPLD and ACQ can be shown using quasinormality, which is a further constraint qualification [3, 7]. Lastly the relation between ACQ and GCQ also follows directly from their definitions. These relations are presented in Figure 3.1.

$$
\begin{tikzcd}
 & \text{MFCQ} \arrow[dr, bend left] & & & \\
\text{LICQ} \arrow[ur, bend left] \arrow[dr, bend right] & & \text{CPLD} \arrow[r] & \text{ACQ} \arrow[r] & \text{GCQ} \\
 & \text{CRCQ} \arrow[ur, bend right] & & &
\end{tikzcd}
$$

Figure 3.1: Implications between constraint qualifications.

3.1.2 Second Order Optimality Conditions

Optimality conditions for (3.1) can also be formulated using second order derivatives, see for example [36, 68, 85]. We repeat two classic second order optimality conditions in this section. For these we will consider a subset of the linearisation cone at a feasible point, the *critical cone*. It consists of all vectors in the linearisation cone which are also possible descent directions of the objective function.

Definition 3.10. Let x^* be a feasible point of (3.1). Then

$$\mathcal{C}_X^{NLP}(x^*) \coloneqq \mathcal{L}_X^{NLP}(x^*) \cap \left\{ d \in \mathbb{R}^n \,:\, \nabla f(x^*)^T d \leq 0 \right\} \tag{3.4}$$

is called *critical cone of X at x^**. A vector $d \in \mathcal{C}_X^{NLP}(x^*)$ is called *critical direction (at x^*)*.

In case that x^* is a KKT point of (3.1), we can give a representation of the critical cone at x^* using its KKT multipliers. The proof of the following lemma is a brief calculation.

Lemma 3.11 (Representation of the Critical Cone with KKT Multipliers)**.** *Let x^* be a KKT point of* (3.1) *with multipliers* (λ^*, μ^*)*. Then*

$$\begin{aligned}
\mathcal{C}_X^{NLP}(x^*) = \{ d \in \mathbb{R}^n \,: & \nabla g_i(x^*)^T d = 0 \quad \forall i \in I_g(x^*) \text{ such that } \lambda_i^* > 0, \\
& \nabla g_i(x^*)^T d \leq 0 \quad \forall i \in I_g(x^*) \text{ such that } \lambda_i^* = 0, \\
& \nabla h_i(x^*)^T d = 0 \quad \forall i = 1, \ldots, p \}.
\end{aligned}$$

Using the critical cone, a necessary second order optimality condition can be formulated for (3.1). The linear independence constraint qualification is required to hold, because the proof relies on an implicit function theorem.

Theorem 3.12 (Second Order Necessary Optimality Condition)**.** *Let x^* be a local minimum of* (3.1) *which satisfies LICQ. Let* (λ^*, μ^*) *be its (due to LICQ unique) KKT multipliers. Then*

$$d^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(x^*) \right) d \geq 0$$

for all $d \in \mathcal{C}_X^{NLP}(x^*)$.

In case a similar second order condition holds, the KKT conditions are sufficient optimality conditions for (3.1).

Theorem 3.13 (Second Order Sufficient Optimality Condition)**.** *Let x^* be a KKT point of* (3.1) *with multipliers* (λ^*, μ^*)*. If*

$$d^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla^2 h_i(x^*) \right) d > 0$$

for all $d \in \mathcal{C}_X^{NLP}(x^*) \setminus \{0\}$, then x^* is a strict local minimum of (3.1).

The inequality in the second order sufficient optimality condition is satisfied for any $d \in \mathbb{R}^n$, if the functions f, g and h are uniformly convex.

3.2 First Order Optimality Conditions for the Complementarity Formulation

In this section we will discuss the reasons why standard constraint qualifications, such as the ones presented in Section 3.1, cannot expected hold for (1.2). We then introduce custom constraint qualifications and custom stationary conditions. These stationary conditions are necessary optimality conditions, provided a custom constraint qualification holds. These were first introduced and discussed in [14, 15]. Furthermore this section contains a few additional results from [13].

3.2.1 Linearisations of the Feasible Set and Constraint Qualifications

First order optimality conditions, such as the KKT conditions, play a central role for the development of numerical methods. However, the KKT conditions are only fulfilled in a local minimum if a constraint qualification holds in that local minimum. Due to the constraints of the complementarity formulation (1.2), constraint qualifications cannot be expected to hold. To clarify this problem, we consider the relation between the linearisation cone and the Bouligand tangent cone for the feasible set of (1.2). We will use a number of index sets for certain active constraints to ease the notation. Let (x^*, y^*) be feasible for (1.2). Like for the nonlinear program (3.1) discussed in Section 3.1, we use the index set

$$I_g(x^*) \coloneqq \{i \in \{1, \dots, m\} \, : \, g_i(x^*) = 0\}$$

for the active inequality constraints. For all indexes such that $x_i^* = 0$ we define

$$I_0(x^*) \coloneqq \{i \in \{1, \dots, n\} \, : \, x_i^* = 0\}.$$

To further distinguish, we also define

$$\begin{aligned}
I_{\pm 0}(x^*, y^*) &\coloneqq \{i \in \{1, \dots, n\} \, : \, x_i^* \neq 0, \, y_i^* = 0\}, \\
I_{00}(x^*, y^*) &\coloneqq \{i \in \{1, \dots, n\} \, : \, x_i^* = 0, \, y_i^* = 0\}, \\
I_{0+}(x^*, y^*) &\coloneqq \{i \in \{1, \dots, n\} \, : \, x_i^* = 0, \, y_i^* \in (0, 1)\}, \\
I_{01}(x^*, y^*) &\coloneqq \{i \in \{1, \dots, n\} \, : \, x_i^* = 0, \, y_i^* = 1\}.
\end{aligned}$$

The above four sets form a partition of $\{1, \dots, n\}$, while the last three sets form a partition of $I_0(x^*)$.
The linearisation cone, compare Definition 3.3, of (1.2) at a feasible point (x^*, y^*) is given by

$$\begin{aligned}\mathcal{L}_Z^{NLP}(x^*,y^*) = \{(d_x,d_y) \in \mathbb{R}^n \times \mathbb{R}^n :\ & \nabla g_i(x^*)^T d_x \le 0 \ \forall i \in I_g(x^*),\\ & \nabla h_i(x^*)^T d_x = 0 \ \forall i = 1,\dots,p,\\ & -e^T d_y \le 0 \text{ if } e^T y^* = n-\kappa,\\ & -e_i^T d_y \le 0 \ \forall i \in I_{\pm 0}(x^*,y^*) \cup I_{00}(x^*,y^*),\\ & e_i^T d_y \le 0 \ \forall i \in I_{01}(x^*,y^*),\\ & y_i^* e_i^T d_x + x_i^* e_i^T d_y = 0 \ \forall i = 1,\dots,n\}.\end{aligned}$$

In our discussion we will use the following equivalent representation of $\mathcal{L}_Z^{NLP}(x^*,y^*)$, see [15, Lemma 3.1]. Its derivation is straightforward. For any feasible point (x^*,y^*) of (1.2) we have

$$\begin{aligned}\mathcal{L}_Z^{NLP}(x^*,y^*) = \{(d_x,d_y) \in \mathbb{R}^n \times \mathbb{R}^n :\ & \nabla g_i(x^*)^T d_x \le 0 \quad \forall i \in I_g(x^*),\\ & \nabla h_i(x^*)^T d_x = 0 \quad \forall i = 1,\dots,p,\\ & e^T d_y \ge 0 \quad \text{if } e^T y^* = n-\kappa,\\ & e_i^T d_y = 0 \quad \forall i \in I_{\pm 0}(x^*,y^*),\\ & e_i^T d_y \ge 0 \quad \forall i \in I_{00}(x^*,y^*),\\ & e_i^T d_y \le 0 \quad \forall i \in I_{01}(x^*,y^*),\\ & e_i^T d_x = 0 \quad \forall i \in I_{0+}(x^*,y^*),\\ & e_i^T d_x = 0 \quad \forall i \in I_{01}(x^*,y^*)\}.\end{aligned} \tag{3.5}$$

If ACQ, or any stronger constraint qualification, holds at (x^*,y^*), recall that this implies $\mathcal{L}_Z^{NLP}(x^*,y^*) = \mathcal{T}_Z(x^*,y^*)$. This relation cannot be expected to hold in the present case. As we will see, the tangent cone of Z at (x^*,y^*) is the union of finitely many cones, which is not convex in general. Yet the linearisation cone of Z at (x^*,y^*) is convex, since by its definition it is polyhedral.

To illuminate this difficulty, we will use an auxiliary problem. Let $J \subseteq I_{00}(x^*,y^*)$. We consider the following nonlinear problem.

$$\begin{aligned}\text{NLP}(J):\quad \min_{x,y}\ f(x) \quad \text{s.t.}\quad & g_i(x) \le 0, && \forall i = 1,\dots,m,\\ & h_i(x) = 0, && \forall i = 1,\dots,p,\\ & e^T y \ge n-\kappa, &&\\ & x_i = 0,\ 0 \le y_i \le 1, && \forall i \in I_{0+}(x^*,y^*) \cup I_{01}(x^*,y^*) \cup J,\\ & y_i = 0, && \forall i \in I_{\pm 0}(x^*,y^*) \cup (I_{00}(x^*,y^*) \setminus J).\end{aligned} \tag{3.6}$$

We will also refer to the above nonlinear program as *piecewise nonlinear program.* It obviously depends on the point (x^*,y^*). Let $Z(J)$ denote the feasible set of NLP(J):

$$\begin{aligned}Z(J) \coloneqq \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n :\ & g_i(x) \le 0, && \forall i = 1,\dots,m,\\ & h_i(x) = 0, && \forall i = 1,\dots,p,\\ & e^T y \ge n-\kappa, &&\\ & x_i = 0,\ 0 \le y_i \le 1, && \forall i \in I_{0+}(x^*,y^*) \cup I_{01}(x^*,y^*) \cup J,\\ & y_i = 0, && \forall i \in I_{\pm 0}(x^*,y^*) \cup (I_{00}(x^*,y^*) \setminus J)\ \}.\end{aligned}$$

Clearly the point (x^*, y^*) is also feasible for NLP(J) and $Z(J) \subseteq Z$ for all $J \subseteq I_{00}(x^*, y^*)$. We proceed to investigate some properties of the piecewise decomposition. Like in the theory on MPCCs, the feasible set Z of (1.2) can be rewritten locally as the union of the feasible sets of all NLP(J). For MPCCs this connection has proven helpful for deriving local properties such as optimality conditions. In Section 3.5 we will use this connection to derive a local error bound for (1.2) . The following statement, first shown in [14], formalises this relation.

Lemma 3.14 ([14, Proposition 4.1])**.** *Let $(x^*, y^*) \in Z$. There exists a neighbourhood $N(x^*, y^*)$ of (x^*, y^*) such that we have*

$$Z \cap N(x^*, y^*) = \left(\bigcup_{J \subseteq I_{00}(x^*, y^*)} Z(J) \right) \cap N(x^*, y^*).$$

Proof. Let $(x, y) \in Z$ and define

$$\hat{J} := \{i \in I_{00}(x^*, y^*) \,:\, x_i = 0,\ y_i > 0\} \cup \{i \in I_{00}(x^*, y^*) \,:\, x_i = 0,\ y_i = 0\}.$$

Then

$$(\hat{J})^C = \{i \in I_{00}(x^*, y^*),\ x_i \neq 0,\ y_i = 0\} \tag{3.7}$$

and, by the above definition, we have $x_i = 0$ for all $i \in \hat{J}$ and $y_i = 0$ for all $i \in (\hat{J})^C$.
Let $i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$. If (x, y) is sufficiently close to (x^*, y^*), i.e. in some neighbourhood $N(x^*, y^*)$ of (x^*, y^*), we have $y_i > 0$. Therefore $x_i = 0$ holds, because (x, y) is feasible for (1.2).
For $i \in I_{\pm 0}(x^*, y^*)$, if (x, y) is sufficiently close to (x^*, y^*), we also have $x_i \neq 0$. Since (x, y) is feasible for (1.2), we have $y_i = 0$. Altogether we have

$$\begin{aligned} x_i = 0,\ y_i \geq 0, \quad & \forall i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*) \cup \hat{J}, \\ y_i = 0, \quad & \forall i \in I_{\pm 0}(x^*, y^*) \cup (\hat{J})^C. \end{aligned}$$

Since the remaining constraints of NLP$(\hat{J})$ are also constraints of (1.2), we have

$$(x, y) \in N(x^*, y^*) \cap Z(\hat{J}) \subseteq N(x^*, y^*) \cap \left(\bigcup_{J \subseteq I_{00}(x^*, y^*)} Z(J) \right).$$

Now let $(x, y) \in N(x^*, y^*) \cap \Big(\bigcup_{J \subseteq I_{00}(x^*, y^*)} Z(J) \Big)$ for some neighbourhood $N(x^*, y^*)$ of (x^*, y^*). Then $(x, y) \in Z(J)$ for some $J \subseteq I_{00}(x^*, y^*)$. Since

$$(I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*) \cup J) \cup \left(I_{\pm 0}(x^*, y^*) \cup J^C \right) = \{1, \ldots, n\}$$

we have $x_i = 0$ or $y_i = 0$ for all $i \in \{1, \ldots, n\}$ and therefore $x_i \cdot y_i = 0$ for all $i \in \{1, \ldots, n\}$. For the same reason we also have $0 \leq y_i \leq 1$ for all $i \in \{1, \ldots, n\}$. Since the remaining constraints of (1.2) are also constraints of NLP(J), we have $(x, y) \in Z$. $\square$

For the analysis of the tangent cone of Z we will use the following relations to the tangent cones of $Z(J)$ $(J \subseteq I_{00}(x^*, y^*))$.

Proposition 3.15 ([15, Proposition 3.2]). *Let* $(x^*, y^*) \in Z$. *Then*

(a) $\mathcal{T}_Z(x^*, y^*) = \bigcup_{J \subseteq I_{00}(x^*,y^*)} \mathcal{T}_{Z(J)}(x^*, y^*)$,

(b) $\mathcal{T}_Z(x^*, y^*)^\circ = \bigcap_{J \subseteq I_{00}(x^*,y^*)} \mathcal{T}_{Z(J)}(x^*, y^*)^\circ$.

Proof. Let $(x^*, y^*) \in Z$.

(a) By Lemma 3.14, there is a neighbourhood $N(x^*, y^*)$ such that

$$Z \cap N(x^*, y^*) = \left(\bigcup_{J \subseteq I_{00}(x^*,y^*)} Z(J) \right) \cap N(x^*, y^*).$$

The tangent cone $\mathcal{T}_Z(x^*, y^*)$ is a local approximation of the feasible set Z, see Definition 3.1. Using this characteristic in the first and third of the following equations, as well as the neighbourhood $N(x^*, y^*)$ in the second equality, we have

$$\begin{aligned}
\mathcal{T}_Z(x^*, y^*) &= \mathcal{T}_{Z \cap N(x^*,y^*)}(x^*, y^*) \\
&= \mathcal{T}_{\left(\bigcup_{J \subseteq I_{00}(x^*,y^*)} Z(J)\right) \cap N(x^*,y^*)}(x^*, y^*) \\
&= \mathcal{T}_{\bigcup_{J \subseteq I_{00}(x^*,y^*)} Z(J)}(x^*, y^*) \\
&= \bigcup_{J \subseteq I_{00}(x^*,y^*)} \mathcal{T}_{Z(J)}(x^*, y^*).
\end{aligned}$$

The last equality holds for unions of finite numbers of sets, as a quick inspection of the tangent cone's definition confirms.

(b) Using part (a), we then have

$$\mathcal{T}_Z(x^*, y^*)^\circ = \left(\bigcup_{J \subseteq I_{00}(x^*,y^*)} \mathcal{T}_{Z(J)}(x^*, y^*) \right)^\circ = \bigcap_{J \subseteq I_{00}(x^*,y^*)} \mathcal{T}_{Z(J)}(x^*, y^*)^\circ,$$

where the second equality follows from [5, Theorem 3.1.9], because there are finitely many $J \subseteq I_{00}(x^*, y^*)$.

□

Lemma 3.15(a) shows that the tangent cone is indeed the union of finitely many cones and therefore not convex in general. Hence it cannot be expected to be equal to the linearisation cone, which is convex. This motivates the introduction of a custom linearisation cone which overcomes this problem. For related constructions see for example [74, 79, 28] for MPCCs and [44] for MPVCs. By intersecting $\mathcal{L}_Z^{NLP}(x^*, y^*)$ with vectors (d_x, d_y) which fulfil the condition

$$(e_i^T d_x)(e_i^T d_y) = 0 \quad \forall i \in I_{00}(x^*, y^*),$$

we now define a linearisation cone that accounts for the complementarity constraint of (1.2).

Definition 3.16. Let $(x^*, y^*) \in Z$. The *CC-linearisation cone* is defined as

$$\begin{aligned}
\mathcal{L}_Z^{CC}(x^*, y^*) \coloneqq \{(d_x, d_y) \in \mathbb{R}^n \times \mathbb{R}^n : \; & \nabla g_i(x^*)^T d_x \leq 0, \quad \forall i \in I_g(x^*), \\
& \nabla h_i(x^*)^T d_x = 0, \quad \forall i = 1, \ldots, p, \\
& e^T d_y \geq 0, \quad \text{if } e^T y^* = n - \kappa, \\
& e_i^T d_y = 0, \quad \forall i \in I_{\pm 0}(x^*, y^*), \\
& e_i^T d_y \geq 0, \quad \forall i \in I_{00}(x^*, y^*), \\
& e_i^T d_y \leq 0, \quad \forall i \in I_{01}(x^*, y^*), \\
& e_i^T d_x = 0, \quad \forall i \in I_{0+}(x^*, y^*), \\
& e_i^T d_x = 0, \quad \forall i \in I_{01}(x^*, y^*), \\
& (e_i^T d_x)(e_i^T d_y) = 0, \quad \forall i \in I_{00}(x^*, y^*)\}.
\end{aligned}$$

As for the tangent cone, we can also obtain a representation of the CC-linearisation cone using the piecewise nonlinear programs (3.6).

Proposition 3.17 ([15, Proposition 3.3])**.** *Let $(x^*, y^*) \in Z$. Then*

(a) $\mathcal{L}_Z^{CC}(x^*, y^*) = \bigcup_{J \subseteq I_{00}(x^*, y^*)} \mathcal{L}_{Z(J)}^{NLP}(x^*, y^*)$,

(b) $\mathcal{L}_Z^{CC}(x^*, y^*)^\circ = \bigcap_{J \subseteq I_{00}(x^*, y^*)} \mathcal{L}_{Z(J)}^{NLP}(x^*, y^*)^\circ$.

Proof. Let $(x^*, y^*) \in Z$.

(a) We have

$$\begin{aligned}
& \bigcup_{J \subseteq I_{00}(x^*, y^*)} \mathcal{L}_{Z(J)}^{NLP}(x^*, y^*) \\
= & \bigcup_{J \subseteq I_{00}(x^*, y^*)} \big\{(d_x, d_y) \in \mathbb{R}^n \times \mathbb{R}^n : \quad \nabla g_i(x^*)^T d_x \leq 0, \quad \forall i \in I_g(x^*), \\
& \qquad \nabla h_i(x^*)^T d_x = 0, \quad \forall i = 1, \ldots, p, \\
& \qquad e^T d_y \geq 0, \quad \text{if } e^T y^* = n - \kappa, \\
& \qquad e_i^T d_x = 0, \quad \forall i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*), \\
& \qquad e_i^T d_x = 0, \quad \forall i \in J, \\
& \qquad e_i^T d_y \geq 0, \quad \forall i \in J, \\
& \qquad e_i^T d_y \leq 0, \quad \forall i \in I_{01}(x^*, y^*), \\
& \qquad e_i^T d_y = 0, \quad \forall i \in I_{\pm 0}(x^*, y^*), \\
& \qquad e_i^T d_y = 0, \quad \forall i \in I_{00}(x^*, y^*) \setminus J\big\}
\end{aligned}$$

$$
\begin{aligned}
=\bigcup\nolimits_{J\subseteq I_{00}(x^*,y^*)} \big\{(d_x,d_y)\in\mathbb{R}^n\times\mathbb{R}^n : \quad \nabla g_i(x^*)^T d_x &\le 0, \quad \forall i\in I_g(x^*),\\
\nabla h_i(x^*)^T d_x &= 0, \quad \forall i=1,\dots,p,\\
e^T d_y &\ge 0, \quad \text{if } e^T y^* = n-\kappa,\\
e_i^T d_x &= 0, \quad \forall i\in I_{0+}(x^*,y^*)\cup I_{01}(x^*,y^*),\\
e_i^T d_y &\ge 0, \quad \forall i\in J,\\
e_i^T d_y &\le 0, \quad \forall i\in I_{01}(x^*,y^*),\\
e_i^T d_y &= 0, \quad \forall i\in I_{\pm 0}(x^*,y^*),\\
(e_i^T d_x)(e_i^T d_y) &= 0, \quad \forall i\in I_{00}(x^*,y^*)\big\}\\
=\mathcal{L}_Z^{CC}(x^*,y^*).&
\end{aligned}
$$

(b) Using part (a), we have

$$
\left(\mathcal{L}_Z^{CC}(x^*,y^*)\right)^\circ = \left(\bigcup_{J\subseteq I_{00}(x^*,y^*)} \mathcal{L}_{Z(J)}^{NLP}(x^*,y^*)\right)^\circ = \bigcap_{J\subseteq I_{00}(x^*,y^*)} \mathcal{L}_{Z(J)}^{NLP}(x^*,y^*)^\circ,
$$

where we apply [5, Theorem 3.1.9] to obtain the last equality.

□

Obviously the CC-linearisation cone is a subset of the linearisation cone. For our further investigations of optimality conditions, the relation $\mathcal{T}_Z(x^*,y^*)\subseteq\mathcal{L}_Z^{CC}(x^*,y^*)$ is desirable. The above results help us to show this relation.

Proposition 3.18 ([15, Proposition 3.4])**.** *Let $(x^*,y^*)\in Z$. Then the following inclusions hold:*

$$
\mathcal{T}_Z(x^*,y^*)\subseteq\mathcal{L}_Z^{CC}(x^*,y^*)\subseteq\mathcal{L}_Z^{NLP}(x^*,y^*). \tag{3.8}
$$

Proof. Let $(x^*,y^*)\in Z$. Since the tangent cone is a subset of the linearisation cone, using Proposition 3.15(a) and Proposition 3.17(a), we obtain

$$
\mathcal{T}_Z(x^*,y^*) = \bigcup_{J\subseteq I_{00}(x^*,y^*)} \mathcal{T}_{Z(J)}(x^*,y^*) \subseteq \bigcup_{J\subseteq I_{00}(x^*,y^*)} \mathcal{L}_{Z(J)}^{NLP}(x^*,y^*) = \mathcal{L}_Z^{CC}(x^*,y^*).
$$

The inclusion $\mathcal{L}_Z^{CC}(x^*,y^*)\subseteq\mathcal{L}_Z^{NLP}(x^*,y^*)$ directly follows from the definition of the CC-linearisation cone. □

Relation (3.8) motivates the following versions of ACQ and GCQ, which were introduced in [15].

Definition 3.19. Let $(x^*,y^*)\in Z$. We say that

(a) *CC-ACQ (Cardinality Constrained Abadie Constraint Qualification)* holds at (x^*,y^*), if $\mathcal{T}_Z(x^*,y^*)=\mathcal{L}_Z^{CC}(x^*,y^*)$,

(b) *CC-GCQ (Cardinality Constrained Guignard Constraint Qualification)* holds at (x^*,y^*), if $\mathcal{T}_Z(x^*,y^*)^\circ=\mathcal{L}_Z^{CC}(x^*,y^*)^\circ$.

It is clear from the above definition that CC-ACQ implies CC-GCQ. Let us consider a basic example of a cardinality constrained optimization problem.

Example 3.20. Let $f : \mathbb{R}^2 \to \mathbb{R}$. We consider the problem

$$\min_{x \in \mathbb{R}^2} f(x) \quad \text{s.t.} \quad \|x\|_0 \leq 1.$$

The complementarity formulation of the above problem is

$$\begin{aligned} \min_{(x,y) \in \mathbb{R}^2 \times \mathbb{R}^2} \ f(x) \quad \text{s.t.} \quad & y_1 + y_2 \geq 1, \\ & 0 \leq y_i \leq 1, \quad \forall i = 1, 2, \\ & x_i y_i = 0, \quad \forall i = 1, 2. \end{aligned} \tag{3.9}$$

Let $x^* = (0, 0)$ and $y^* = (0, 1)$. The point (x^*, y^*) is feasible for (3.9). To check that ACQ holds at (x^*, y^*) we compute the cones $\mathcal{L}_Z^{NLP}(x^*, y^*)$ and $\mathcal{T}_Z(x^*, y^*)$. We have

$$I_{00}(x^*, y^*) = \{1\}, \ I_{01}(x^*, y^*) = \{2\} \text{ and } I_{0+}(x^*, y^*) = I_{\pm 0}(x^*, y^*) = \emptyset.$$

Using representation (3.5) of the linearisation cone at (x^*, y^*), we have

$$\mathcal{L}_Z^{NLP}(x^*, y^*) = \{(d_x, d_y) \in \mathbb{R}^2 \times \mathbb{R}^2 \, : \, d_{x_2} = 0, \ d_{y_1} \geq 0, \ d_{y_2} \leq 0, \ d_{y_1} + d_{y_2} \geq 0\}.$$

For the tangent cone we use Proposition 3.15(a). The feasible sets of the piecewise nonlinear programs NLP(J) for $J \in \mathcal{P}(I_{00}(x^* y^*)) = \{\emptyset, \{1\}\}$ are given by

$$\begin{aligned} Z(\emptyset) &= \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \, : \, y_1 + y_2 \geq 1, \ x_2 = 0, \ 0 \leq y_2 \leq 1, \ y_1 = 0\}, \\ Z(\{1\}) &= \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \, : \, y_1 + y_2 \geq 1, \ x_1 = 0, \ x_2 = 0, \ 0 \leq y_1 \leq 1, \ 0 \leq y_2 \leq 1\}. \end{aligned}$$

Since the above sets are polyhedral we have

$$\begin{aligned} \mathcal{T}_{Z(\emptyset)}(x^*, y^*) &= \mathcal{L}_{Z(\emptyset)}^{NLP}(x^*, y^*) \\ &= \{(d_x, d_y) \in \mathbb{R}^2 \times \mathbb{R}^2 \, : \, d_{x_2} = 0, \ d_{y_1} = d_{y_2} = 0\}, \\ \mathcal{T}_{Z(\{1\})}(x^*, y^*) &= \mathcal{L}_{Z(\{1\})}^{NLP}(x^*, y^*) \\ &= \{(d_x, d_y) \in \mathbb{R}^2 \times \mathbb{R}^2 \, : \, d_{x_1} = d_{x_2} = 0, \ d_{y_1} \geq 0, \ d_{y_2} \leq 0, \ d_{y_1} + d_{y_2} \geq 0\}. \end{aligned}$$

By Proposition 3.15(a) we have

$$\mathcal{T}_Z(x^*, y^*) = \mathcal{T}_{Z(\emptyset)}(x^*, y^*) \cup \mathcal{T}_{Z(\{1\})}(x^*, y^*).$$

Let $(d_x, d_y) = ((1, 0), (1, 0))$. Then $(d_x, d_y) \in \mathcal{L}_Z^{NLP}(x^*, y^*) \setminus \mathcal{T}_Z(x^*, y^*)$, hence ACQ does not hold in (x^*, y^*). However, by Proposition 3.17(a) and Proposition 3.15(a) we have

$$\begin{aligned} \mathcal{L}_Z^{CC}(x^*, y^*) &= \bigcup_{J \subseteq I_{00}(x^*, y^*)} \mathcal{L}_{Z(J)}^{NLP}(x^*, y^*) \\ &= \mathcal{L}_{Z(\emptyset)}^{NLP}(x^*, y^*) \cup \mathcal{L}_{Z(\{1\})}^{NLP}(x^*, y^*) \\ &= \mathcal{T}_{Z(\emptyset)}(x^*, y^*) \cup \mathcal{T}_{Z(\{1\})}(x^*, y^*) \\ &= \bigcup_{J \subseteq I_{00}(x^*, y^*)} \mathcal{T}_{Z(J)}(x^*, y^*) \\ &= \mathcal{T}_Z(x^*, y^*). \end{aligned}$$

Consequently CC-ACQ holds at (x^*, y^*).

Even in this simple example the linearisation cone does not coincide with the tangent cone, hence ACQ and every stronger constraint qualification is violated. The following result is central for the relation between GCQ and CC-GCQ: Although the CC-linearisation cone is a subset of the linearisation cone, their polar cones coincide.

Theorem 3.21 ([15, Theorem 3.7])**.** *Let $(x^*, y^*) \in Z$. Then $\mathcal{L}_Z^{NLP}(x^*, y^*)^\circ = \mathcal{L}_Z^{CC}(x^*, y^*)^\circ$.*

From the above theorem we directly obtain the following result.

Corollary 3.22 ([15, Corollary 3.8])**.** *Let $(x^*, y^*) \in Z$. Then CC-GCQ holds at (x^*, y^*) if and only if GCQ holds at (x^*, y^*).*

The above result is quite different from the theory on MPCCs. In that setting an LICQ-type constraint qualification, which is stronger than CC-GCQ, implies (standard) GCQ, see [29]. We will comment on the implications of Corollary 3.22 again in the following section, after we introduce stationary conditions for (1.2).
In [14, 15] further constraint qualifications for (1.2) were introduced using the so-called *tightened nonlinear program.* Like the piecewise nonlinear program, this problem is defined with respect to a given feasible point (x^*, y^*) of (1.2), yet depends only on the part x^*:

$$\begin{aligned} \text{TNLP}(x^*): \quad \min\nolimits_{x,y}\ f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0, \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0, \quad \forall i = 1, \dots, p, \\ & x_i = 0, \qquad \forall i \in I_0(x^*). \end{aligned} \tag{3.10}$$

The above problem is a restriction of the original cardinality constrained problem (1.1) to the support of x^*. The feasible set of TNLP(J) is always a subset of the feasible set of (1.1), hence the name tightened nonlinear program. A similar program is also used in the theory for mathematical programs with complementarity constraints to define constraint qualifications and optimality conditions [81, 87]. We say that the *cardinality constrained linear independence constraint qualification* (CC-LICQ) holds at (x^*, y^*) for (1.2), if LICQ holds at x^* for TNLP(x^*). Analogously we can define constraint qualifications corresponding to MFCQ, CRCQ and CPLD. We give a formal definition.

Definition 3.23. Let (x^*, y^*) be feasible for (1.2). We say that (x^*, y^*) or x^* satisfies

(a) *CC-LICQ (Cardinality Constrained - Linear Independence Constraint Qualification)* if the gradients

$$\nabla g_i(x^*),\ i \in I_g(x^*),\ \nabla h_i(x^*),\ i = 1, \dots, p,\ e_i,\ i \in I_0(x^*),$$

are linearly independent.

(b) *CC-MFCQ (Cardinality Constrained - Mangasarian-Fromovitz Constraint Qualification)* if the gradients

$$\left\{\nabla g_i(x^*),\ i \in I_g(x^*)\right\} \quad \text{and} \quad \left\{\nabla h_i(x^*),\ i = 1, \dots, p,\ e_i,\ i \in I_0(x^*)\right\}$$

are positively linearly independent.

(c) *CC-CRCQ (Cardinality Constrained - Constant Rank Constraint Qualification)* if for any subset $I_1 \subseteq I_g(x^*)$, $I_2 \subseteq \{1, \dots, p\}$ and $I_3 \subseteq I_0(x^*)$ such that the gradients

$$\nabla g_i(x),\ i \in I_1,\ \nabla h_i(x),\ i \in I_2,\ e_i,\ i \in I_3,$$

are linearly dependent in $x = x^*$, they remain linearly dependent in a neighbourhood of x^*.

(d) *CC-CPLD (Cardinality Constrained - Constant Positive Linear Dependence Constraint Qualification)* if for any subset $I_1 \subseteq I_g(x^*)$, $I_2 \subseteq \{1, \dots, p\}$ and $I_3 \subseteq I_0(x^*)$ such that the gradients

$$\big\{\nabla g_i(x),\ i \in I_1\big\} \quad \text{and} \quad \big\{\nabla h_i(x),\ i \in I_2,\ e_i,\ i \in I_3\big\}$$

are positively linearly dependent at $x = x^*$, they remain linearly dependent in a neighbourhood of x^*.

We synonymously say that a CC-constraint qualification *holds at* x^*, if x^* satisfies the CC-constraint qualification in question. The above constraint qualifications solely depend on the part x^*. They can thus be regarded as constraint qualifications for the original problem (1.1). Because they are defined using constraint qualifications for the tightened nonlinear program, the relations between the above constraint qualifications correspond to the relations between the constraint qualifications for a standard nonlinear program. The CC-LICQ implies CC-MFCQ and CC-CRCQ. Both CC-MFCQ and CC-CRCQ imply CC-CPLD. It also follows from their definition, that CC-ACQ implies CC-GCQ. For the proof that CC-CPLD implies CC-ACQ, we will use the piecewise nonlinear program once more. The implications between the CC-constraint qualifications are illustrated in Figure 3.2. The CC-CPLD constraint qualification implies the CPLD constraint qualifications for NLP(J) for all $J \subseteq I_{00}(x^*, y^*)$, as stated by the following lemma.

Lemma 3.24 ([15, Lemma 3.12])**.** *Let* (x^*, y^*) *be a feasible point of* (1.2) *and let CC-CPLD be fulfilled in* (x^*, y^*)*. Then CPLD holds for NLP(J) in* (x^*, y^*) *for any* $J \subseteq I_{00}(x^*, y^*)$*.*

Proof. Let $J \subseteq I_{00}(x^*, y^*)$ and

$$I_1 \subseteq I_g(x^*), \quad I_2 \subseteq \begin{cases} \{0\}, & \text{if } \sum_{i=1}^n y_i^* = n - \kappa, \\ \emptyset, & \text{otherwise,} \end{cases}, \quad I_3 \subseteq I_{01}(x^*, y^*), \quad I_4 \subseteq J,$$
$$I_5 \subseteq \{1, \dots, p\}, \quad I_6 \subseteq I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*) \cup J, \quad I_7 \subseteq I_{\pm 0}(x^*, y^*) \cup J^C.$$

Let the gradients of the relevant active constraints of (3.6)

$$\begin{aligned} &\left\{\begin{pmatrix}\nabla g_i(x)\\0\end{pmatrix}, i \in I_1, \begin{pmatrix}0\\-e\end{pmatrix}, i \in I_2, \begin{pmatrix}0\\e_i\end{pmatrix}, i \in I_3, \begin{pmatrix}0\\-e_i\end{pmatrix}, i \in I_4\right\} \\ \text{and} \quad &\left\{\begin{pmatrix}\nabla h_i(x)\\0\end{pmatrix}, i \in I_5, \begin{pmatrix}e_i\\0\end{pmatrix}, i \in I_6, \begin{pmatrix}0\\e_i\end{pmatrix}, i \in I_7\right\}\end{aligned} \tag{3.11}$$

be positively linearly dependent in $(x, y) = (x^*, y^*)$. Then there exist coefficients

$$0 \neq (a, b_0, c, d, \alpha, \beta, \gamma) \in \mathbb{R}^{|I_1|} \times \mathbb{R}^{|I_2|} \times \mathbb{R}^{|I_3|} \times \mathbb{R}^{|I_4|} \times \mathbb{R}^{|I_5|} \times \mathbb{R}^{|I_6|} \times \mathbb{R}^{|I_7|}$$

such that

$$\sum_{i\in I_1} a_i \nabla g_i(x) + \sum_{i\in I_5} \alpha_i \nabla h_i(x) + \sum_{i\in I_6} \beta_i e_i = 0, \tag{3.12}$$

$$b_0(-e) + \sum_{i\in I_3} c_i e_i + \sum_{i\in I_4} d_i(-e_i) + \sum_{i\in I_7} \gamma_i e_i = 0, \tag{3.13}$$

$$a_i \geq 0 \ \forall i \in I_1,\ b_0 \geq 0,\ c_i \geq 0 \ \forall i \in I_3,\ d_i \geq 0 \ \forall i \in I_4, \tag{3.14}$$

holds for $(x, y) = (x^*, y^*)$. In case we have either $a_i > 0$ for some $i \in I_1$, $\alpha_i \neq 0$ for some $i \in I_5$ or $\beta_i \neq 0$ for some $i \in I_6$, it follows from (3.12) and (3.14), that the vectors

$$\{\nabla g_i(x), i \in I_1\} \quad \text{and} \quad \{\nabla h_i(x), i \in I_5, e_i, i \in I_6\} \tag{3.15}$$

are positively linearly dependent in $x = x^*$. Since $I_6 \subseteq I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*) \cup J_2 \subseteq I_0(x^*)$ and CC-CPLD holds in (x^*, y^*), the vectors (3.15) are linearly dependent in a neighbourhood of x^*. Consequently the vectors (3.11) are linearly dependent for all (x, y) in a neighbourhood of (x^*, y^*).
In case we have either $b_0 > 0$, $c_i > 0$ for some $i \in I_3$, $d_i > 0$ for some $i \in I_4$ or $\gamma_i \neq 0$ for some $i \in I_7$, it follows from (3.13), that the vectors

$$\{-e, i \in I_2,\ e_i, i \in I_3,\ -e_i, i \in I_4,\ e_i, i \in I_6\}$$

are linearly dependent. Hence also in this case, the vectors (3.11) are linearly dependent for all (x, y) in a neighbourhood of (x^*, y^*). □

The next lemma states an implication from ACQ for the tightened nonlinear programs to CC-ACQ.

Lemma 3.25 ([15, Lemma 3.9])**.** *Let $(x^*, y^*) \in Z$ and $J \subseteq I_{00}(x^*, y^*)$. Let (x^*, y^*) satisfy ACQ for $Z(J)$ for each piecewise nonlinear program NLP(J). Then CC-ACQ holds at (x^*, y^*).*

With this implication we can prove the relation between CC-CPLD and CC-ACQ.

Theorem 3.26 ([15, Theorem 3.13])**.** *Let $(x^*, y^*) \in Z$ satisfy CC-CPLD. Then CC-ACQ holds at (x^*, y^*).*

Proof. Since (x^*, y^*) satisfies CC-CPLD, by Lemma 3.24, the CPLD constraint qualification holds at (x^*, y^*) for $Z(J)$ for each $J \subseteq I_{00}(x^*, y^*)$. Hence ACQ holds at (x^*, y^*) for each $Z(J)$. From Lemma 3.25 it follows that CC-ACQ holds at (x^*, y^*). □

Analogously to Lemma 3.24, we can also show a relation between CC-MFCQ for (1.2) and MFCQ for the piecewise nonlinear problems.

Lemma 3.27. *Let (x^*, y^*) be a feasible point of (1.2). If CC-MFCQ holds in (x^*, y^*), then MFCQ for NLP(J) in (x^*, y^*) holds for all $J \subseteq I_{00}(x^*, y^*)$ such that $I_{0+(x^*,y^*)} \cup J \neq \emptyset$ or $e^T y^* > n - \kappa$.*

Proof. To show that MFCQ for NLP(J) holds at (x^*, y^*), we have to show that the gradients

$$\begin{pmatrix} \nabla g_i(x^*) \\ 0 \end{pmatrix} (i \in I_g), \quad \begin{pmatrix} 0 \\ -e \end{pmatrix} (\text{if } e^T y^* = n - \kappa), \quad \begin{pmatrix} 0 \\ -e_i \end{pmatrix} (i \in J), \quad \begin{pmatrix} 0 \\ e_i \end{pmatrix} (i \in I_{01})$$

$$\text{and} \quad \begin{pmatrix} \nabla h_i(x^*) \\ 0 \end{pmatrix} (i = 1, \ldots, p), \quad \begin{pmatrix} e_i \\ 0 \end{pmatrix} (i \in I_{0+} \cup I_{01} \cup J), \quad \begin{pmatrix} 0 \\ e_i \end{pmatrix} (i \in I_{\pm 0} \cup J^C)$$

$$
\begin{tikzcd}
 & \text{CC-MFCQ} \arrow[dr, bend left] & & & \\
\text{CC-LICQ} \arrow[ur, bend left] \arrow[dr, bend right] & & \text{CC-CPLD} \arrow[r] & \text{CC-ACQ} \arrow[r] & \text{CC-GCQ} \\
 & \text{CC-CRCQ} \arrow[ur, bend right] & & &
\end{tikzcd}
$$

Figure 3.2: Implications between CC-constraint qualifications.

are positively linearly independent. Since all gradients have either an x- or a y-part but never on both, we can consider them separately. The gradients with the x-parts are positively linearly independent due to CC-MFCQ. And the gradients with the y-parts are positively linearly independent if $e^T y^* > n - \kappa$ or $I_{01}(x^*, y^*) \cup I_{\pm 0}(x^*, y^*) \cup J^C \neq \{1, \dots, n\}$, where the latter is equivalent to $J \cup I_{0+}(x^*, y^*) \neq \emptyset$. $\square$

We conclude this section with two results on CC-constraint qualifications. Since they depend only on the x-part of a given feasible point (x, y), CC-MFCQ as well as CC-CPLD still hold in a neighbourhood regarding the x-part. This fact will be used in the convergence analysis of a Scholtes-type regularisation method in Chapter 4.

Lemma 3.28. *Let $(x^*, y^*) \in Z$ satisfy CC-MFCQ. Then there exists an $r > 0$ such that every $(x, y) \in Z$ with $x \in B_r(x^*)$ also satisfies CC-MFCQ.*

Proof. Assume the claim is false. Then there exists a sequence $(x^k, y^k)_{k \in \mathbb{N}} \subseteq Z$ with $x^k \to x^*$ as $k \to \infty$ such that CC-MFCQ does not hold in (x^k, y^k) for every $k \in \mathbb{N}$. Hence there are coefficients $(a^k, b^k, c^k) \in (\mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n) \setminus \{0\}$ such that

$$
\sum_{i=1}^{m} a_i^k \nabla g_i(x^k) + \sum_{i=1}^{p} b_i^k \nabla h_i(x^k) + \sum_{i=1}^{n} c_i^k e_i = 0, \tag{3.16}
$$
$$
a_i^k \geq 0 \;\forall i = 1, \dots, m, \quad a_i^k = 0 \;\forall i \notin I_g(x^k), \quad c_i^k = 0 \;\forall i \notin I_0(x^k)
$$

for all $k \in \mathbb{N}$. This means the relevant gradients are positively linearly dependent in (x^k, y^k) for all $k \in \mathbb{N}$. Because $(a^k, b^k, c^k) \neq 0$ we can assume without loss of generality $\|(a^k, b^k, c^k)\| = 1$ for all $k \in \mathbb{N}$. Then the sequence is convergent (at least on a subsequence). Let

$$
\lim_{k \to \infty}(a^k, b^k, c^k) = (a, b, c).
$$

We have $(a, b, c) \neq 0$ and $a_i \geq 0$ for all i. Since g is continuous, we have $g_i(x^k) < 0$ for all $i \notin I_g(x^*)$ and all sufficiently large $k \in \mathbb{N}$. Hence $I_g(x^k) \subseteq I_g(x^*)$ and analogously $I_0(x^k) \subseteq I_0(x^*)$ for all sufficiently large $k \in \mathbb{N}$. Using the same argument, we further have $a_i = 0$ for all $i \notin I_g(x^*)$ and $c_i = 0$ for all $i \notin I_0(x^*)$. Since h and g are continuous, it follows from (3.16) that

$$
\sum_{i \in I_g(x^*)} a_i \nabla g_i(x^*) + \sum_{i=1}^{p} b_i \nabla h_i(x^*) + \sum_{i \in I_0(x^*)} c_i e_i = 0
$$

for $k \to \infty$. This is a contradiction to CC-MFCQ in (x^*, y^*). $\square$

The next lemma is the analogous result for CC-CPLD, which will also play a role in Chapter 4.

Lemma 3.29. *Let $(x^*, y^*) \in Z$ satisfy CC-CPLD. Then there exists an $r > 0$ such that every $(x, y) \in Z$ with $x \in B_r(x^*)$ also satisfies CC-CPLD.*

Proof. Assume the claim is false. Then there exists a sequence $(x^k)_{k\in\mathbb{N}}$ converging to x^* such that CC-CPLD is violated in x^k for every $k \in \mathbb{N}$. Thus there exist index sets $I_1^k \subseteq I_g(x^k)$, $I_2^k \subseteq \{1, \dots, p\}$ and $I_3^k \subseteq I_0(x^k)$ for every k such that

$$\{\nabla g_i(x^k),\ i \in I_1^k\} \quad \text{and} \quad \{\nabla h_i(x^k), i \in I_2^k,\ e_i, i \in I_3^k\},$$

are positively linearly dependent, and for every k there is a sequence $(x^{k,j})_{j\in\mathbb{N}}$ such that the gradients

$$\nabla g_i(x^{k,j}),\ i \in I_1^k,\ \nabla h_i(x^{k,j}),\ i \in I_2^k,\ e_i,\ i \in I_3^k, \tag{3.17}$$

are linearly independent for every $j \in \mathbb{N}$. Since $x^k \to x^*$ $(k \to \infty)$ we have $I_1^k \subseteq I_g(x^*)$, $I_2^k \subseteq \{1, \dots, p\}$ and $I_3^k \subseteq I_0(x^*)$ for sufficiently large $k \in \mathbb{N}$. This can be argued analogously to the proof of Lemma 3.28. Moreover, since the index sets $I_g(x^*)$, $\{1, \dots, p\}$ and $I_0(x^*)$ are finite, we can find $I_1 \subseteq I_g(x^*)$, $I_2 \subseteq \{1, \dots, p\}$ and $I_3 \subseteq I_0(x^*)$ such that $I_1^k = I_1$, $I_2^k = I_2$ and $I_3^k = I_3$ for all k (at least on a subsequence, which we also denote by $(x^k)_k$). Thus for every k the gradients

$$\{\nabla g_i(x^k),\ i \in I_1\} \quad \text{and} \quad \{\nabla h_i(x^k), i \in I_2,\ e_i, i \in I_3\}, \tag{3.18}$$

are positively linearly dependent. For every k there exists an index $j(k) \in \mathbb{N}$ such that $\|x^{k,j(k)} - x^k\| \leq \|x^k - x^*\|$. Consequently we have

$$\|x^{k,j(k)} - x^*\| \leq \|x^{j(k),k} - x^k\| + \|x^k - x^*\| \leq 2 \cdot \|x^k - x^*\| \longrightarrow 0 \quad (k \to \infty).$$

Because $x^k \to x^*$ $(k \to \infty)$ and (3.18) holds, the gradients

$$\{\nabla g_i(x^*),\ i \in I_1\} \quad \text{and} \quad \{\nabla h_i(x^*), i \in I_2,\ e_i, i \in I_3\},$$

are positively linearly dependent (this can be shown analogously to the proof of Lemma 3.28). Setting $(\xi^k)_{k\in\mathbb{N}} := (x^{k,j(k)})_{k\in\mathbb{N}}$ we have $\xi^k \to x^*$. From (3.17) it follows that the gradients

$$\nabla g_i(\xi^k),\ i \in I_1,\ \nabla h_i(\xi^k), i \in I_2,\ e_i, i \in I_3,$$

are linearly independent for every k. This yields a contradiction to CC-CPLD in x^*. □

As we will see in the following section, first order optimality conditions hold under the CC-constraint qualification which we discussed in this section.

3.2.2 Stationarity and Optimality Conditions

In this section we will establish first order optimality conditions for the complementarity formulation which hold under CC-constraint qualifications. We will use the following two stationary concepts introduced in [14].

Definition 3.30. A feasible point $(x^*, y^*) \in Z$ of (1.2) is called

(a) *M-stationary* (M = Mordukhovich), if there exist multipliers $(\lambda^*, \mu^*, \gamma^*) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$ such that

$$\begin{aligned}\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla h_i(x^*) + \sum_{i=1}^n \gamma_i^* e_i &= 0,\\ \lambda_i^* \geq 0, \quad \lambda_i^* \cdot g_i(x^*) &= 0, \quad \forall i = 1, \ldots, m,\\ \gamma_i^* &= 0, \quad \forall i \in I_{\pm 0}(x^*, y^*).\end{aligned}$$

(b) *S-stationary* (S = Strong), if it is M-stationary and additionally $\gamma_i^* = 0 \; \forall i \in I_{00}(x^*, y^*)$.

The name M-stationarity is motivated from stationarity conditions for mathematical programs with complementarity constraints. The corresponding condition for those is derived using the *limiting normal cone*, which is also known as the Mordukhovich normal cone. We will look into the relation to mathematical programs with complementarity constraints at the end of this section.
Every S-stationary point is also an M-stationary point, as one can see from the above definition. This implication is strict, i.e. not every M-stationary point is an S-stationary point (see for instance [15, Example 5.5]). M-stationarity only depends on the variable x^*. Therefore we also call x^* itself M-stationary, meaning that there exists a vector y such that (x^*, y) is feasible for (1.2), and x^* satisfies the definition of M-stationarity. While serving as first order necessary optimality conditions, the conditions S- and M-stationarity will also play a role in the convergence analysis in Chapter 4. The stronger condition, S-stationarity, is equivalent to the KKT conditions for (1.2).

Proposition 3.31 ([14, Proposition 4.7])**.** *Let $(x^*, y^*) \in Z$. Then (x^*, y^*) is a KKT point of* (1.2) *if and only if it is an S-stationarity point of* (1.2).

Proof. Let $(x^*, y^*) \in Z$ be a KKT point of (1.2) with multipliers $(\lambda, \mu, \tilde{\gamma}, \delta, \nu^-, \nu^+) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ such the KKT conditions are fulfilled:

$$\begin{aligned}\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \mu_i \nabla h_i(x^*) + \sum_{i=1}^n \tilde{\gamma}_i y_i^* e_i &= 0,\\ -\delta e + \sum_{i=1}^n \nu_i^-(-e_i) + \sum_{i=1}^n \nu_i^+ e_i + \sum_{i=1}^n \tilde{\gamma}_i x_i^* e_i &= 0,\\ \lambda_i \geq 0, \quad \lambda_i \cdot g_i(x^*) &= 0, \quad \forall i = 1, \ldots, m,\\ \delta \geq 0, \quad \delta \cdot (e^T y^* - n + \kappa) &= 0,\\ \nu_i^-, \nu_i^+ \geq 0, \quad \nu_i^+ \cdot y_i^* = 0, \quad \nu_i^+ \cdot (y_i^* - 1) &= 0, \quad \forall i = 1, \ldots, n.\end{aligned} \tag{3.19}$$

Let $\gamma_i := \tilde{\gamma}_i \cdot y_i^*$ for all $i = 1, \ldots, n$. Then $\gamma_i = 0$ for all $i \in I_{00}(x^*, y^*) \cup I_{\pm 0}(x^*, y^*)$ and we have

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \mu_i \nabla h_i(x^*) + \sum_{i=1}^n \gamma_i e_i = 0.$$

Thus (x^*, y^*) is an S-stationary point of (1.2) with multipliers (λ, μ, γ).

Conversely, let (x^*, y^*) be an S-stationary point of (1.2) with multipliers (λ, μ, γ). Let

$$\tilde{\gamma}_i := \begin{cases} \frac{\gamma_i}{y_i^*}, & \text{if } i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*), \\ 0, & \text{if } i \in I_{00}(x^*, y^*) \cup I_{\pm 0}(x^*, y^*). \end{cases}$$

Since (x^*, y^*) is S-stationary, we then have

$$\begin{aligned} &\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i=1}^{n} \tilde{\gamma}_i y_i^* e_i \\ =&\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i=1}^{n} \gamma_i e_i = 0. \end{aligned}$$

Setting $(\delta, \nu_i^-, \nu_i^+) = (0, 0, 0)$, it is easy to see that the point (x^*, y^*) along with the multipliers $(\lambda, \mu, \tilde{\gamma}, \nu^-, \nu^+)$ satisfies all KKT conditions (3.19). □

Under CC-GCQ, or any stronger CC-constraint qualification, a local minimum of (1.2) is an S-stationary point. We will use the previous proposition to prove this.

Theorem 3.32 ([15, Theorem 4.2])**.** *Let (x^*, y^*) be a local minimum of* (1.2) *such that CC-GCQ holds at (x^*, y^*). Then (x^*, y^*) is an S-stationary point.*

Proof. Since CC-GCQ holds at (x^*, y^*), by Corollary 3.22, GCQ also holds at (x^*, y^*). Then by Theorem 3.9 the point (x^*, y^*) is a KKT point. Thus, by Proposition 3.31, the point (x^*, y^*) is also S-stationary. □

Theorem 3.32 is quite different from the theory on MPCCs for which an LICQ-type condition is required for the corresponding result. For (1.2) the KKT conditions already hold under CC-GCQ or any stronger CC-constraint qualification. This is a consequence of Theorem 3.21. It states that the polar cone of the linearisation cone and the polar cone of the CC-linearisation cone in a given feasible point of (1.2) are identical. Hence GCQ is implied by CC-GCQ or any stronger CC-constraint qualification.
In case the constraint functions g and h are linear, a local minimum of (1.2) is always an S-stationary point, as stated by the following corollary.

Corollary 3.33 ([15, Corollary 4.3])**.** *Let the functions g and h be linear and let (x^*, y^*) be a local minimum of* (1.2)*. Then (x^*, y^*) is S-stationary.*

Proof. Because g and h are linear, all constraints of the piecewise nonlinear program NLP(J), for each $J \subseteq I_{00}(x^*, y^*)$, are in fact linear. Hence ACQ holds at (x^*, y^*) for $Z(J)$ for each $J \subseteq I_{00}(x^*, y^*)$, see for example [36]. By Lemma 3.25 CC-ACQ thus holds at (x^*, y^*). Since CC-ACQ implies CC-GCQ, the statement then follows from Theorem 3.32. □

The CC-LICQ guarantees that a local minimum (x^*, y^*) of (1.2) is S-stationary and it is not hard to see that the corresponding multipliers are unique. In case x^* is also a local minimum of the original problem (1.1), a similar result can be obtained for all points (x^*, y) feasible for (1.2).

Proposition 3.34. *Let x^* be a local minimum of* (1.1) *satisfying CC-LICQ. Then every point $(x^*, y) \in Z$ is S-stationary. The corresponding multipliers $(\lambda^*, \mu^*, \gamma^*) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$ are unique and independent from y. In case $\|x^*\|_0 < \kappa$, we additionally have $\gamma^* = 0$.*

Proof. Since x^* is a local minimum of (1.1), for all y such that $(x^*, y) \in Z$ the point (x^*, y) is a local minimum of (1.2) and thus due to CC-LICQ an S-stationary point, see [15, Theorem 4.2]. Hence there exist S-stationary multipliers $(\lambda^*, \mu^*, \gamma^*)$ with $\lambda_i^* = 0$ for all $i \notin I_g(x^*)$, $\gamma_i^* = 0$ for all $i \notin I_0(x^*)$ and

$$\nabla f(x^*) + \sum_{i \in I_g(x^*)} \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla h_i(x^*) + \sum_{i \in I_0(x^*)} \gamma_i^* e_i = 0.$$

Due to CC-LICQ this equation has at most one solution $(\lambda_{I_g}^*, \mu^*, \gamma_{I_0}^*)$ and thus the multipliers $(\lambda^*, \mu^*, \gamma^*)$ are unique and independent from y.
Let $\|x^*\|_0 < \kappa$. It remains to show that in this case $\gamma^* = 0$. For all $i \notin I_0(x^*)$ this is guaranteed by the definition of S-stationarity. For every $j \in I_0(x^*)$ we can define

$$y_i = \begin{cases} 0 & \text{if } i \in \operatorname{supp}(x^*) \cup \{j\}, \\ 1 & \text{else.} \end{cases} \tag{3.20}$$

Because $|I_0(x^*)| > n - \kappa$ the point (x^*, y) is feasible for (1.2) and thus a local minimum and S-stationary point of (1.2). The S-stationarity conditions then imply $\gamma_j^* = 0$. Since the multipliers $(\lambda^*, \mu^*, \gamma^*)$ are unique and independent from y and the same argument holds for all $j \in I_0(x^*)$, we have shown $\gamma^* = 0$. □

In case (1.1) is convex, except for the cardinality constraint of course, S-stationarity is a sufficient first order optimality condition.

Theorem 3.35 ([15, Theorem 4.4])**.** *Assume that the functions f and g are convex and the function h is affine linear. Let (x^*, y^*) be an S-stationary point of* (1.2)*. Then (x^*, y^*) is a local minimum of* (1.2)*.*

Proof. Let $(x, y) \in Z$ and let (λ, μ, γ) be the multipliers of (x^*, y^*). Using the S-stationarity of (x^*, y^*), we obtain

$$\begin{aligned} f(x) &\geq f(x^*) + \nabla f(x^*)^T (x - x^*) \\ &= f(x^*) - \sum_{i \in I_g} \underbrace{\lambda_i}_{\geq 0} \underbrace{\nabla g_i(x^*)^T (x - x^*)}_{\leq g_i(x) - g_i(x^*) = g_i(x) \leq 0} - \sum_{i=1}^{p} \mu_i \underbrace{\nabla h_i(x^*)(x - x^*)}_{= h_i(x) - h_i(x^*) = 0} - \sum_{i \in I_{0+} \cup I_{01}} \gamma_i e_i^T (x - x^*) \\ &\geq f(x^*) - \sum_{i \in I_{0+} \cup I_{01}} \gamma_i e_i^T (x - x^*) \\ &= f(x^*) - \sum_{i \in I_{0+} \cup I_{01}} \gamma_i x_i. \end{aligned}$$

For the first inequality we use the convexity of f. The second inequality is obtained by using the convexity of g as well as the fact that h is affine linear. If (x, y) is sufficiently close to (x^*, y^*) we have $y_i > 0$ for all $i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$. Since (x, y) is feasible, this implies $x_i = 0$ for all $i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$. Consequently there exists a radius $r > 0$ such that

$$f(x) \geq f(x^*) \quad \forall (x, y) \in Z \cap B_r(x^*, y^*).$$

□

To conclude our discussion of first order optimality conditions we will have a look at similar concepts for mathematical programs with complementarity constraints. As mentioned earlier, the complementarity formulation is very similar to problems from this class. The following comparisons have been established in [15].
Let $N, M, P, Q \in \mathbb{N}$, $f : \mathbb{R}^N \to \mathbb{R}$, $g : \mathbb{R}^N \to \mathbb{R}^M$, $h : \mathbb{R}^N \to \mathbb{R}^P$, $G : \mathbb{R}^N \to \mathbb{R}^Q$ and $H : \mathbb{R}^N \to \mathbb{R}^Q$. A *mathematical program with complementarity constraints (MPCC)* is of the form

$$\begin{aligned} \min_{x\in\mathbb{R}^n} f(x) \quad \text{s.t.} \quad & g_i(x) \le 0, && \forall i = 1, \dots, M, \\ & h_i(x) = 0, && \forall i = 1, \dots, P, \\ & G_i(x) \ge 0,\ H_i(x) \ge 0,\ H_i(x)G_i(x) = 0, && \forall i = 1, \dots, Q. \end{aligned} \tag{3.21}$$

Like for the previously introduced optimality conditions, we assume all involved functions to be continuously differentiable. If there are additional constraints $x \ge 0$ present in (1.2), we can consider it as an MPCC of the form (3.21). In this special case the complementarity formulation is given by

$$\begin{aligned} \min_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n} f(x) \quad \text{s.t.} \quad & g_i(x) \le 0, && \forall i = 1, \dots, m, \\ & -x_i \le 0, && \forall i = 1, \dots, n, \\ & h_i(x) = 0, && \forall i = 1, \dots, p, \\ & e^T y \ge n - \kappa, && \\ & 0 \le y_i \le 1, && \forall i = 1, \dots, n, \\ & x_i \cdot y_i = 0, && \forall i = 1, \dots, n. \end{aligned} \tag{3.22}$$

There are two ways to consider (3.22): The first way is to regard (3.22) as an instance of (1.2) by redefining the inequality constraints as $(g(x), -x) \le 0$. The second way is to regard (3.22) as an MPCC of the form (3.21). To this end we add the constraints $e^T y \ge n - \kappa$ and $y_i \le 0$, $i = 1, \dots, n$, to the inequality constraints g in (3.21). By setting $G_i(x, y) := x_i$, $H_i(x, y) := y_i$, $i = 1, \dots, n$, we account for the constraints $x_i \ge 0$, $y_i \ge 0$ and $x_i \cdot y_i = 0$ (we have $N = 2 \cdot n$, $M = m + 1 + n$, $P = p$ and $Q = n$ in the setting of (3.21)).
Firstly, let us consider (3.22) as an special instance of (1.2). The following conditions for M-stationarity and S-stationarity hold for this problem.

Lemma 3.36 ([15, Lemma 5.2])**.** *Let* (x^*, y^*) *be feasible for* (3.22)*. Then* (x^*, y^*) *is*

(a) S-stationary (in the sense of Definition 3.30) if and only if there exist multipliers $(\lambda, \mu, \gamma) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$ *such that*

$$\begin{aligned} \nabla f(x^*) + \sum_{i\in I_g(x^*)} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i\in I_0(x^*)} \gamma_i e_i = 0, \\ \lambda_i \ge 0, \quad \forall i \in I_g(x^*), \\ \gamma_i \le 0, \quad \forall i \in I_{00}(x^*, y^*). \end{aligned}$$

(b) M-stationary (in the sense of Definition 3.30) if and only if there exist multipliers $(\lambda, \mu, \gamma) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$ such that

$$\nabla f(x^*) + \sum_{i \in I_g(x^*)} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i \in I_0(x^*)} \gamma_i e_i = 0,$$
$$\lambda_i \geq 0, \quad \forall i \in I_g(x^*).$$

Although we have additional constraints $x \geq 0$ in (3.22) the conditions for M-stationary points are exactly the same as for the general form of the complementarity formulation. For S-stationarity there are additional multipliers γ_i for $i \in I_{00}(x^*, y^*)$ with a sign restriction present.
Now let us consider (3.22) as an MPCC. We repeat four stationary conditions for MPCCs, see for example [81, 87].

Definition 3.37. Let x^* be feasible for (3.21). Then x^* is called

(a) *W-stationary* (W = weakly), if there are multipliers $(\lambda, \mu, \gamma, \nu) \in \mathbb{R}^M \times \mathbb{R}^P \times \mathbb{R}^Q \times \mathbb{R}^Q$ such that

$$\begin{aligned} \nabla f(x^*) &+ \sum_{i: g_i(x^*)=0} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{P} \mu_i \nabla h_i(x^*) \\ &- \sum_{i: G_i(x^*)=0} \gamma_i \nabla G_i(x^*) - \sum_{i: H_i(x^*)=0} \nu_i \nabla H_i(x^*) = 0, \\ &\lambda_i \geq 0, \qquad \forall i \in I_g(x^*), \end{aligned}$$

(b) *C-stationary* (C = Clarke), if it is W-stationary and additionally for all i with $G_i(x^*) = H_i(x^*) = 0$ we have

$$\gamma_i \nu_i \geq 0,$$

(c) *M-stationary* (M = Mordukhovich), if it is W-stationary and additionally for all i with $G_i(x^*) = H_i(x^*) = 0$ we have

$$\big(\gamma_i \geq 0 \text{ and } \nu_i \geq 0\big) \quad \text{or} \quad \gamma_i \nu_i = 0,$$

(d) *S-stationary* (S = strongly), if it is W-stationary and additionally for all i with $G_i(x^*) = H_i(x^*) = 0$ we have

$$\gamma_i \geq 0 \text{ and } \nu_i \geq 0.$$

The strongest of the above stationarity conditions is S-stationarity. Like for the S-stationary conditions for (1.2) (in the sense of Definition 3.30), it can be shown that S-stationarity for (3.21) (in the sense of Definition 3.37) is equivalent to the KKT conditions. The relations of the above conditions are as follows: S-stationarity implies M-stationarity, M-stationarity implies C-stationarity, which in turn implies W-stationarity. The following lemma states the relations between M- and S-stationarity of the complementarity formulation with nonnegative constraints on x and their MPCC counterparts.

Lemma 3.38 ([15, Lemma 5.3]). *Let* (x^*, y^*) *be feasible for* (3.22). *Then* (x^*, y^*) *is S-stationary (M-stationary) in the sense of Definition 3.30 if and only if it is S-stationary (M-stationary) in the sense of Definition* (3.37).

Moreover, for a feasible point (x^*, y^*) of (3.22) we have that

$$(x^*, y^*) \text{ is M-stationary} \quad \Leftrightarrow \quad (x^*, y^*) \text{ is C-stationary} \quad \Leftrightarrow \quad (x^*, y^*) \text{ is W-stationary}$$

in the sense of Definition 3.37 (see [15, Remark 5.4]). This means the (MPCC) concepts M-, C- and W-stationarity coincide for the special case (3.22) with M-stationarity in the sense of Definition 3.30.
For S-stationarity in the sense of Definition 3.37 to be a necessary optimality condition we have to assume an LICQ-type constraint qualification, which is a strong contrast to Theorem 3.32. The names of the stationarity conditions M- and S-stationarity for (1.2) (in the sense of Definition 3.30) are motivated by their counterparts for MPCCs. In both settings M-stationarity can be deduced using the limiting normal cone, which is also known as *Mordukhovich* normal cone.

3.3 Second Order Optimality Conditions for the Complementarity Formulation

In this section we will expand the set of optimality conditions for the complementarity-type formulation (1.2) of cardinality constrained optimization problems (1.1) to second order optimality conditions. The results in this section were established in joint work with Alexandra Schwartz [13]. We begin in Section 3.3.1 with a second order necessary condition which holds under CC-CRCQ. In Section 3.3.2 we then derive a second order sufficient condition for S-stationary points. It captures the lack of curvature of the objective function f regarding the auxiliary variable y. These results relate to the classic second order optimality conditions (Theorem 3.12 and Theorem 3.13). For M-stationary points, we show that they are unique with respect to the variable x under CC-CPLD and a second order condition in Section 3.3.3.
To formulate second order optimality conditions, we will use a subset of the CC-linearisation cone $\mathcal{L}_Z^{CC}((x^*, y^*))$ at a given point $(x^*, y^*) \in Z$. The *CC-critical cone*, see also [43, 59, 68] for related constructions, is the set of all potentially feasible descent directions.

Definition 3.39. Let $(x^*, y^*) \in Z$. The *CC-critical cone* of Z at (x^*, y^*) is defined by

$$\mathcal{C}_Z^{CC}(x^*, y^*) := \mathcal{L}_Z^{CC}(x^*, y^*) \ \cap \ \left\{(d_x, d_y) \in \mathbb{R}^n \times \mathbb{R}^n \ : \ \nabla f(x^*)^T d_x \leq 0\right\}.$$

A vector $d \in \mathcal{C}_Z^{CC}(x^*, y^*)$ is called *critical direction* (*at* (x^*, y^*)).

Later, we are particularly interested in the directions d_x only. It is straightforward to see that for any $(x^*, y^*) \in Z$ we have

$$\begin{aligned}
&\{d_x \in \mathbb{R}^n \ : \ \exists d_y \in \mathbb{R}^n \ : \ (d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)\} \\
= &\{d_x \in \mathbb{R}^n \ : \ \nabla g_i(x^*)^T d_x \leq 0, \quad \forall i \in I_g(x^*), \\
&\qquad\qquad\quad \nabla h_i(x^*)^T d_x = 0, \quad \forall i = 1, \dots, p, \\
&\qquad\qquad\quad e_i^T d_x = 0, \quad \forall i \in I_{01}(x^*, y^*) \cup I_{0+}(x^*, y^*)\}.
\end{aligned}$$

In case $\|x^*\|_0 < \kappa$, this set still depends on the chosen value of y^*. Thus for a given x^* we also consider the union over all y^* such that $(x^*, y^*) \in Z$:

$$\mathcal{L}_{\mathcal{X}}(x^*) := \{d_x \in \mathbb{R}^n \ : \ \exists y^* \in \mathbb{R}^n, d_y \in \mathbb{R}^n \ : (x^*, y^*) \in Z, \ (d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)\}.$$

The following representation of $\mathcal{L}_{\mathcal{X}}(x^*)$ will be helpful.

Lemma 3.40. *Let x^* be feasible for* (1.1). *Then we have*

$$\mathcal{L}_{\mathcal{X}}(x^*) = \{d_x \in \mathbb{R}^n \quad : \quad \begin{array}{ll} \nabla g_i(x^*)^T d_x \leq 0 & \forall i \in I_g(x^*), \\ \nabla h_i(x^*)^T d_x = 0 & \forall i = 1, \dots, p, \\ |\{i \in I_0(x^*) \ : \ (d_x)_i = 0\}| \geq n - \kappa\}. & \end{array} \tag{3.23}$$

Proof. "$\subseteq$" Let $d_x \in \mathcal{L}_{\mathcal{X}}(x^*)$. Hence there are vectors y^* and d_y such that $(x^*, y^*) \in Z$ and $(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)$. We have $e_i^T d_x = 0$ for all $i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*) \subseteq I_0(x^*)$ and thus

$$|\{i \in I_0(x^*) \ : \ (d_x)_i = 0\}| \geq |I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)| \geq e^T y^* \geq n - \kappa.$$

"$\supseteq$" Let

$$d_x \in \{d_x \in \mathbb{R}^n \quad : \quad \begin{array}{ll} \nabla g_i(x^*)^T d_x \leq 0 & \forall i \in I_g(x^*), \\ \nabla h_i(x^*)^T d_x = 0 & \forall i = 1, \dots, p, \\ |\{i \in I_0(x^*) \ : \ (d_x)_i = 0\}| \geq n - \kappa\}. & \end{array}$$

Setting

$$y_i^* := \begin{cases} 0, & \text{if } x_i^* \neq 0, \\ 1, & \text{if } x_i^* = 0 \text{ and } (d_x)_i = 0, \\ 0, & \text{if } x_i^* = 0 \text{ and } (d_x)_i \neq 0, \end{cases}$$

for $i = 1, \dots, n$, we have $(x^*, y^*) \in Z$ and $I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*) = I_{01}(x^*, y^*)$. Let $i \in I_{01}(x^*, y^*)$ be arbitrary. Then we have $x_i^* = 0$, $y_i^* = 1$ and hence by definition of y^* also $(d_x)_i = 0$. Thus we have $e_i^T d_x = 0$ for all $i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$. Set $d_y := 0$. Then all remaining constraints in the definition of $\mathcal{L}_Z^{CC}(x^*, y^*)$ are clearly satisfied as well. Consequently $(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)$ and hence $d_x \in \mathcal{L}_{\mathcal{X}}(x^*)$.

□

The set $\mathcal{L}_{\mathcal{X}}(x^*)$ can be seen as a linearisation cone of the original feasible set $\mathcal{X}$. Analogously to the CC-critical cone, we define

$$\mathcal{C}_{\mathcal{X}}(x^*) \quad := \quad \{d_x \in \mathcal{L}_{\mathcal{X}}(x^*) \ : \ \nabla f(x^*)^T d_x \leq 0\}.$$

If (x^*, y^*) is an S-stationary point of (1.2), we can give a description of $\mathcal{C}_Z^{CC}(x^*, y^*)$ that does not use the gradient of the objective function but instead the multipliers of (x^*, y^*). For some multipliers $\lambda^* \in \mathbb{R}_+^m$ corresponding to the inequality constraints $g(x^*) \leq 0$ we define the index sets

$$\begin{aligned} I_{g+}(x^*, \lambda^*) \quad &:= \quad \{i \in I_g(x^*) \ : \ \lambda_i^* > 0\}, \\ I_{g0}(x^*, \lambda^*) \quad &:= \quad \{i \in I_g(x^*) \ : \ \lambda_i^* = 0\}. \end{aligned}$$

With these index sets, the following proposition gives a characterisation of the CC-critical cone at an S-stationary point. It corresponds to Lemma 3.11 for the critical cone for standard nonlinear optimization problems.

Proposition 3.41. *Let (x^*, y^*) be an S-stationary point of (1.2) with multipliers $(\lambda^*, \mu^*, \gamma^*)$. Then we have*

$$\mathcal{C}_Z^{CC}(x^*, y^*) = \{(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*) \,:\, \nabla g_i(x^*)^T d_x = 0 \ \forall i \in I_{g+}(x^*, \lambda^*)\}.$$

Proof. Let $(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)$ be arbitrary. It suffices to show the equivalence

$$\nabla f(x^*)^T d_x \leq 0 \quad \Leftrightarrow \quad \nabla g_i(x^*) d_x = 0 \quad \forall i \in I_{g+}(x^*, \lambda^*).$$

Since (x^*, y^*) is S-stationary with multipliers $(\lambda^*, \mu^*, \gamma^*)$ we know $\lambda^* \geq 0$ and

$$\nabla f(x^*) = -\sum_{i \in I_g} \lambda_i^* \nabla g_i(x^*) - \sum_{i=1}^{p} \mu_i^* \nabla h_i(x^*) - \sum_{i \in I_{0+} \cup I_{01}} \gamma_i^* e_i.$$

Taking into account $(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)$, we obtain

$$\begin{aligned}
& 0 \geq \nabla f(x^*)^T d_x \\
\Leftrightarrow \quad & 0 \geq -\sum_{i \in I_g} \lambda_i^* \nabla g_i(x^*)^T d_x - \sum_{i=1}^{p} \mu_i^* \nabla h_i(x^*)^T d_x - \sum_{i \in I_{0+} \cup I_{01}} \gamma_i^* e_i^T d_x \\
\Leftrightarrow \quad & 0 \geq -\sum_{i \in I_g} \lambda_i^* \nabla g_i(x^*)^T d_x \\
\Leftrightarrow \quad & 0 = \nabla g_i(x^*)^T d_x \quad \forall i \in I_{g+}(x^*, \lambda^*),
\end{aligned}$$

since $\lambda_i^* \geq 0$ for all $i = 1, \dots, m$. □

Note that the alternative representation of the CC-critical cone from Proposition 3.41 does not necessarily hold for M-stationary points. The reason is that for an M-stationary point (x^*, y^*) with multipliers $(\lambda^*, \mu^*, \gamma^*)$ and a vector $(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)$ the equation $\gamma_i^* e_i^T d_x = 0$ does not necessary hold for $i \in I_{00}(x^*, y^*)$.

3.3.1 Second Order Necessary Optimality Condition

Our next goal is to derive a second order necessary optimality condition for local minima of the continuous reformulation (1.2). Similar results known from classical nonlinear optimization, see for example [68], and from mathematical programs with vanishing constraints, see [43], are based on an implicit function theorem and therefore would require CC-LICQ. The same line of argument is possible in our case, yet requires CC-LICQ. Instead, we follow an idea from [40] and only require the weaker CC-CRCQ.

Theorem 3.42 (Second Order Necessary Optimality Condition)**.** *Let f, g, h be twice continuously differentiable and (x^*, y^*) be a local minimum of (1.2) satisfying CC-CRCQ. Then (x^*, y^*) is S-stationary and for every triple (λ, μ, γ) of S-stationary multipliers we have*

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla^2 h_i(x^*) \right) d_x \geq 0$$

for all $(d_x, d_y)^T \in \mathcal{C}_Z^{CC}(x^*, y^*)$.

Proof. It follows from Theorem 3.32 that the point (x^*, y^*) is an S-stationary point of (1.2). Let (λ, μ, γ) be arbitrary S-stationary multipliers. We define the Lagrange function $\ell : \mathbb{R}^n \to \mathbb{R}$ as

$$x \mapsto \ell(x) := f(x) + \lambda^T g(x) + \mu^T h(x) + \gamma^T x$$

for $x \in \mathbb{R}^n$. Because the functions f, g and h are twice continuously differentiable the function ℓ is also twice continuously differentiable and

$$\nabla \ell(x) = \nabla f(x) + \nabla g(x)\lambda + \nabla h(x)\mu + \gamma,$$
$$\nabla^2 \ell(x) = \nabla^2 f(x) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x) + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x),$$

for $x \in \mathbb{R}^n$. Because (x^*, y^*) is S-stationary we have

$$\ell(x^*) = f(x^*) \quad \text{and} \quad \nabla \ell(x^*, \lambda, \gamma, \mu) = 0.$$

We will use the auxiliary set

$$\tilde{Z} := Z \cap \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid g_i(x) = 0 \ \forall i \in I_{g+}(x^*, \lambda)\}.$$

Let $(d_x, d_y) \in \mathcal{T}_{\tilde{Z}}(x^*, y^*)$. Then there exist sequences $(x^k, y^k) \subseteq \tilde{Z}$, $t^k \geq 0$ such that $(x^k, y^k) \to (x^*, y^*)$, $t^k \to 0$ and

$$t^k((x^k, y^k) - (x^*, y^*)) \longrightarrow (d_x, d_y) \quad \text{for } k \to \infty.$$

Since (x^*, y^*) is a local minimum of (1.2) we have

$$\ell(x^k) = f(x^k) + \lambda^T g(x^k) + \mu^T h(x^k) + \gamma^T x^k = f(x^k) \geq f(x^*) \tag{3.24}$$

for all $k \in \mathbb{N}$ sufficiently large. The second equation above holds, because $(x^k, y^k) \in \tilde{Z}$ and (λ, μ, γ) are S-stationary multipliers of (x^*, y^*). For all $k \in \mathbb{N}$ Taylor's theorem provides us with a $\xi^k \in [x^*, x^k]$ such that

$$\begin{aligned}
\ell(x^k) &= \ell(x^*) + \underbrace{\nabla \ell(x^*)^T}_{=0}(x^k - x^*) + \frac{1}{2}(x^* - x^k)^T \nabla^2 \ell(\xi^k)(x^k - x^*) \\
&= f(x^*) + \frac{1}{2}(x^* - x^k)^T \nabla^2 \ell(\xi^k)(x^k - x^*).
\end{aligned}$$

Using (3.24) it follows that

$$(x^* - x^k)^T \nabla^2 \ell(\xi^k)(x^k - x^*) \geq 0.$$

Multiplying the above inequality by $(t^k)^2 \geq 0$ and taking the limit $k \to \infty$ we have

$$d_x^T \nabla^2 \ell(x^*) d_x \geq 0 \quad \forall (d_x, d_y) \in \mathcal{T}_{\tilde{Z}}(x^*, y^*), \tag{3.25}$$

because we chose $(d_x, d_y) \in \mathcal{T}_{\tilde{Z}}(x^*, y^*)$ arbitrary.
Now let us consider the sets Z and $\tilde{Z}$ again. For the set $\tilde{Z}$ we have the additional equality constraints $g_i(x) = 0$ for $i \in I_{g+}(x^*, \lambda)$. Since $I_{g+}(x^*, \lambda) \subseteq I_g(x^*)$ the constraints that are active in (x^*, y^*) are the same for both sets. Thus it is easy to see that CC-CRCQ for Z

in (x^*, y^*) is equivalent to CC-CRCQ in (x^*, y^*) for $\tilde{Z}$. Consequently, using that CC-CRCQ implies CC-ACQ for $\tilde{Z}$, we have $\mathcal{T}_{\tilde{Z}}(x^*, y^*) = \mathcal{L}^{CC}_{\tilde{Z}}(x^*, y^*)$.
Now let $(d_x, d_y) \in \mathcal{C}^{CC}_Z(x^*, y^*)$. Using Proposition 3.41 we have

$$\begin{aligned}
\mathcal{C}^{CC}_Z(x^*, y^*) &= \{(d_x, d_y) \in \mathbb{R}^n \times \mathbb{R}^n \mid \nabla g_i(x^*)^T d_x = 0 \ \forall i \in I_{g+}(x^*, \lambda)\} \\
&= \mathcal{L}^{CC}_{\tilde{Z}}(x^*, y^*) = \mathcal{T}_{\tilde{Z}}(x^*, y^*).
\end{aligned}$$

From (3.25) we thus obtain

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i \nabla^2 h_i(x^*) \right) d_x \geq 0$$

for all $(d_x, d_y) \in \mathcal{C}^{CC}_Z(x^*, y^*)$. □

If x^* is a local minimum of (1.1) satisfying CC-LICQ (and thus CC-CRCQ), we know that every feasible point $(x^*, y) \in Z$ is a local minimum and thus an S-stationary point of (1.2). By Proposition 3.34 all S-stationary points (x^*, y) share a unique multiplier $(\lambda^*, \mu^*, \gamma^*)$. Thus, as a corollary we immediately recover the second order necessary optimality condition from [71, Theorem 4.1]:

Corollary 3.43. *Let f, g and h be twice continuously differentiable, x^* be a local minimum of* (1.1) *satisfying CC-LICQ, and $(\lambda^*, \mu^*, \gamma^*)$ be the unique S-stationary multiplier for all $(x^*, y) \in Z$. Then*

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(x^*) \right) d_x \geq 0$$

for all $d_x \in \mathcal{C}_{\mathcal{X}}(x^*)$.

3.3.2 Second Order Sufficient Optimality Condition

In this section we state a second order sufficient optimality condition for (1.2). We begin by introducing a condition for S-stationary points that can be used to identify which S-stationary points are local minima of (1.2).

Definition 3.44. Let f, g, h be twice continuously differentiable, and $(x^*, y^*) \in Z$ be an S-stationary point of (1.2). If for all directions $(d_x, d_y) \in \mathcal{C}^{CC}_Z(x^*, y^*)$ with $d_x \neq 0$ there exists <u>at least one</u> S-stationary multiplier $(\lambda^*, \mu^*, \gamma^*)$ such that

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(x^*) \right) d_x > 0, \tag{3.26}$$

then we say that the *Cardinality Constrained Second Order Sufficient Optimality Condition* (*CC-SOSC*) holds in (x^*, y^*).

Note that in this condition we do not have to check all $(d_x, d_y) \in \mathcal{C}^{CC}_Z(x^*, y^*)$ with $(d_x, d_y) \neq 0$ but only those with $d_x \neq 0$. For directions with $d_x = 0$ condition (3.26) obviously cannot be satisfied. But since the objective function f does only depend on x, the directions $d_x \neq 0$ are the important ones.

From standard nonlinear optimization, we know that a second order sufficiency condition combined with a KKT point yields a strict local minimum. However, since the objective function here does not depend on y, we cannot expect to obtain a strict local minimum with respect to both variables unless y is locally fixed. For this reason, we have to work with the concept of a strict local minimum with respect to x.

Definition 3.45. We say that a feasible point (x^*, y^*) of (1.2) is a *strict local minimum with respect to x of* (1.2), if there exists a radius $r > 0$ such that

$$f(x^*) < f(x) \quad \forall (x, y) \in B_r(x^*, y^*) \cap \{(x, y) \in Z \, : \, x \neq x^*\}.$$

Note that a strict local minimum (x^*, y^*) with respect to x is always a local minimum with respect to both variables since for all $(x, y) \in B_r(x^*, y^*)$ either $x = x^*$ and thus $f(x) = f(x^*)$ or $x \neq x^*$ and thus $f(x) > f(x^*)$.
The following theorem shows that CC-SOSC is indeed a sufficient condition for an S-stationary point to be a local minimum of the reformulation (1.2). For the proof we adapt a line of argument from [43].

Theorem 3.46 (Second Order Sufficient Optimality Condition)**.** *Let f, g, h be twice continuously differentiable and (x^*, y^*) be an S-stationary point of* (1.2) *satisfying CC-SOSC. Then (x^*, y^*) is a strict local minimum with respect to x of* (1.2)*.*

Proof. Assume that the claim is false. Then we can find a sequence $\left(x^k, y^k\right)_k \subseteq Z$ with $(x^k, y^k) \to (x^*, y^*)$ $(k \to \infty)$ and $x^k \neq x^*$ such that $f(x^k) \leq f(x^*)$ for all $k \in \mathbb{N}$. We deduce a contradiction to (3.26) from those properties. To this end define the directions $d^k = (d_x^k, d_y^k)$ by

$$d_x^k := \frac{x^k - x^*}{\|x^k - x^*\|}, \qquad d_y^k := \frac{y^k - y^*}{\|(x^k, y^k) - (x^*, y^*)\|}$$

for all $k \in \mathbb{N}$. We have $\|d_x^k\| = 1$ and $\|d_y^k\| \leq 1$ for all $k \in \mathbb{N}$, i.e. the sequences are bounded. Thus, we can assume without loss of generality that $(d^k)_k$ converges to some direction $d = (d_x, d_y)$. Because $\|d_x^k\| = 1$ for all $k \in \mathbb{N}$ we know $d_x \neq 0$.
We proceed to show that d is a critical direction. To do so, we exploit the fact that (x^k, y^k) are feasible for all $k \in \mathbb{N}$ and converging to (x^*, y^*).
For all $k \in \mathbb{N}$, by the mean value theorem, there is a $\xi^k \in [x^k, x^*]$ such that

$$\nabla g_i(\xi^k)^T (x^k - x^*) = g_i(x^k) - g_i(x^*) \leq 0 \quad \forall i \in I_g(x^*).$$

Dividing the above inequality by $\|x^k - x^*\|$ and letting $k \to \infty$, we obtain $\nabla g_i(x^*)^T d_x \leq 0$ for all $i \in I_g(x^*)$, since ∇g_i is continuous. Analogously we can show $\nabla h_i(x^*)^T d_x = 0$ for all $i = 1, \ldots, p$.
If $e^T y^* = n - \kappa$, we obtain for all $k \in \mathbb{N}$

$$e^T (y^k - y^*) = e^T y^k - (n - \kappa) \geq 0 \quad \Longrightarrow \quad e^T d_y \geq 0.$$

For $i \in I_{\pm 0}(x^*, y^*)$ we have $x_i^k \neq 0$ and thus $y_i^k = 0$ for sufficiently large k. Hence also

$$e_i^T (y^k - y^*) = y_i^k - y_i^* = 0 \quad \Longrightarrow \quad e_i^T d_y = 0.$$

For $i \in I_{00}(x^*, y^*)$ we have

$$e_i^T(y^k - y^*) = y_i^k - y_i^* = y_i^k \geq 0 \quad \Longrightarrow \quad e_i^T d_y \geq 0.$$

For $i \in I_{01}(x^*, y^*)$ we have

$$e_i^T(y^k - y^*) = y_i^k - y_i^* = y_i^k - 1 \leq 0 \quad \Longrightarrow \quad e_i^T d_y \leq 0.$$

For $i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$ we have $y_i^k > 0$ and thus $x_i^k = 0$ for k sufficiently large. Hence also

$$e_i^T(x^k - x^*) = x_i^k - x_i^* = 0 \quad \Longrightarrow \quad e_i^T d_x = 0.$$

For $i \in I_{00}(x^*, y^*)$ we have

$$\begin{aligned}
(e_i^T d_x)(e_i^T d_y) &= \lim_{k\to\infty} \left(\frac{e_i^T(x^k - x^*)}{\|x^k - x^*\|}\right)\left(\frac{e_i^T(y^k - y^*)}{\|(x^k, y^k) - (x^*, y^*)\|}\right) \\
&= \lim_{k\to\infty} \left(\frac{x_i^k - x_i^*}{\|x^k - x^*\|}\right)\left(\frac{y_i^k - y_i^*}{\|(x^k, y^k) - (x^*, y^*)\|}\right) \\
&= \lim_{k\to\infty} \frac{x_i^k \cdot y_i^k}{\|x^k - x^*\|\|(x^k, y^k) - (x^*, y^*)\|} = 0.
\end{aligned}$$

We thus have shown $d \in \mathcal{L}_Z^{CC}(x^*, y^*)$. For all $k \in \mathbb{N}$, applying the mean value theorem to the objective function, we find a $\zeta^k \in [x^k, x^*]$ with

$$\nabla f(\zeta^k)^T(x^k - x^*) = f(x^k) - f(x^*) \leq 0 \quad \Longrightarrow \quad \nabla f(x^*)^T d_x \leq 0.$$

Hence $d \in \mathcal{L}_Z^{CC}(x^*, y^*) \cap \{(d_x, d_y) \in \mathbb{R}^n \times \mathbb{R}^n \,:\, \nabla f(x^*)^T d \leq 0\} = \mathcal{C}_Z^{CC}(x^*, y^*)$.
Now it remains to show that for all S-stationary multipliers $(\lambda^*, \mu^*, \gamma^*)$ the direction $(d_x, d_y) \in \mathcal{C}_Z^{CC}(x^*, y^*)$ violates (3.26). To this end fix an arbitrary S-stationary multiplier $(\lambda^*, \mu^*, \gamma^*)$ and define the twice continuously differentiable function $\ell : \mathbb{R}^n \to \mathbb{R}$ by

$$x \mapsto \ell(x) \coloneqq f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) + \sum_{i=1}^p \mu_i^* h_i(x) + \sum_{i=1}^n \gamma_i^* x_i.$$

The Hessian of ℓ at x^* is the Hessian in (3.26). Using the S-stationarity of (x^*, y^*) with the multipliers $(\lambda^*, \mu^*, \gamma^*)$, we know $\ell(x^*) = f(x^*)$ and $\nabla \ell(x^*) = 0$. For sufficiently large $k \in \mathbb{N}$ we thus obtain

$$\begin{aligned}
\ell(x^*) = f(x^*) &\geq f(x^k) \\
&\geq f(x^k) + \sum_{i=1}^m \lambda_i^* g_i(x^k) + \sum_{i=1}^p \mu_i^* h_i(x^k) + \sum_{i=1}^n \gamma_i^* x_i^k = \ell(x^k).
\end{aligned} \tag{3.27}$$

For the second inequality above we use the feasibility of (x^k, y^k) and thus add only non-positive sums. The last sum is zero due to the fact that $y_i^k > 0$ for all $i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$ and sufficiently large $k \in \mathbb{N}$ and thus $x_i^k = 0$. (Note that this argument does not work if (x^*, y^*) is only M-stationary.) For each $k \in \mathbb{N}$ Taylor's theorem provides us with a $\xi^k \in [x^k, x^*]$ for which the equality

$$\ell(x^k) = \ell(x^*) + \nabla \ell(x^*)^T(x^k - x^*) + \frac{1}{2}(x^k - x^*)^T \nabla^2 \ell(\xi^k)(x^k - x^*)$$

holds. From (3.27) we know $\ell(x^k) - \ell(x^*) \leq 0$. Together with $\nabla\ell(x^*) = 0$ and the above equality, we therefore have

$$\begin{aligned} &(x^k - x^*)^T \left(\nabla^2 f(\xi^k) + \sum_{i=1}^n \lambda_i^* \nabla^2 g_i(\xi^k) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(\xi^k)\right)(x^k - x^*) \\ &= (x^k - x^*)^T \nabla^2 \ell(\xi^k)(x^k - x^*) \;=\; 2(\ell(x^k) - \ell(x^*)) \;\leq\; 0 \end{aligned}$$

for sufficiently large $k \in \mathbb{N}$. Dividing by $\|x^k - x^*\|^2$ and letting k tend to infinity this yields a contradiction to the assumption (3.26) due to $d_x \neq 0$. □

In the previous result we have seen that CC-SOSC is a sufficient condition for a local minimum. However, contrary to the corresponding result in nonlinear programming, it guarantees a strict local minimum only with respect to changes in the x-variable. Such a behaviour was to be expected, since the objective function f does not depend on the variable y. Thus no point (x, y), at which we can change y without changing x, can be a strict local minimum.
This effect can also be observed in the CC-SOSC: The matrix in (3.26) depends only on the x-variable. Thus the expression only depends on the d_x-part of a critical direction $d = (d_x, d_y)$, whereas the set of critical directions depends on both x and y. For this reason, we have to exclude all critical directions $d = (d_x, d_y)$ with $d_x \neq 0$ from the strict inequality (3.26). In contrast, in the SOSC from nonlinear programming and similar results for MPCCs, see for example [68] and [59], only the vector $d = (d_x, d_y) = (0, 0)$ is excluded from the condition.
At first glance the CC-SOSC looks very similar to the (NLP-) Second Order Sufficient Optimality Condition. Let (x^*, y^*) be an S-stationary point with multipliers (λ, μ, γ) and consider the second order condition from Theorem 3.13 applied to (1.2). For $(d_x, d_y) \in \mathbb{R}^n \times \mathbb{R}^n$ we have

$$\begin{aligned} &(d_x, d_y)^T \begin{pmatrix} \nabla^2 f(x^*) + \sum\limits_{i=1}^m \lambda_i \nabla^2 g_i(x^*) + \sum\limits_{i=1}^p \mu_i \nabla^2 h_i(x^*) & \operatorname{diag}(\gamma) \\ \operatorname{diag}(\gamma) & 0 \end{pmatrix} \begin{pmatrix} d_x \\ d_y \end{pmatrix} \\ &= d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i \nabla^2 h_i(x^*)\right) d_x + 2 \cdot \sum_{i=1}^n \gamma_i (d_x)_i (d_y)_i. \end{aligned}$$

For all $(d_x, d_y) \in \mathcal{L}_Z^{NLP}(x^*, y^*)$ we have

$$2 \cdot \sum_{i=1}^n \gamma_i (d_x)_i (d_y)_i = 2 \cdot \sum_{i \in I_{00}(x^*, y^*)} \gamma_i (d_x)_i (d_y)_i.$$

This follows directly from the representation (3.5) of $\mathcal{L}_Z^{NLP}(x^*, y^*)$. If we have $(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)$, the above term actually vanishes. The Hessian matrices in Theorem 3.13 and Theorem 3.46 indeed coincide in this case. However, the (NLP-) Second Order Sufficient Optimality Condition Theorem 3.13 demands critical directions to be in $\mathcal{L}_Z^{NLP}(x^*, y^*)$, which is a superset of $\mathcal{L}_Z^{CC}(x^*, y^*)$.
Besides using a smaller set of critical directions, the exclusion of all directions with $d_x = 0$ is sensible. Whenever the cardinality constraint is inactive in a local minimum, one can find critical directions with $d_x = 0$, $d_y \neq 0$. For these directions the strict inequality (3.26) cannot hold. Thus, excluding only the vector $(d_x, d_y) = (0, 0)$ from (3.26) would lead to a condition which is rarely satisfied. The following example illustrates this point.

Example 3.47. Consider the cardinality constrained optimization problem

$$\min_{x \in \mathbb{R}^2} f(x) = x_1^2 + x_2^2 \quad \text{s.t.} \quad g(x) = x_1^2 + x_2^2 - 1 \leq 0, \ \|x\|_0 \leq 1.$$

The point $x^* = (0,0)$ is the strict (global) minimum of this problem and together with $y^* = (1,0)$ the point (x^*, y^*) is a solution for the continuous reformulation

$$\begin{aligned} \min_{(x,y)} f(x) = x_1^2 + x_2^2 \quad \text{s.t.} \quad & g(x) = x_1^2 + x_2^2 - 1 \leq 0, \\ & 0 \leq y_i \leq 1, \ x_i \cdot y_i = 0 \qquad \forall i = 1,2, \\ & y_1 + y_2 \geq 1. \end{aligned}$$

We have $I_g(x^*) = \emptyset$, $I_0(x^*) = \{1,2\}$ and thus CC-LICQ is fulfilled in (x^*, y^*). Hence (x^*, y^*) is an S-stationary point. Due to CC-LICQ, the corresponding S-stationary multipliers (λ^*, γ^*) are unique. We have $\lambda^* = \gamma^* = 0$. Since $\nabla f(x^*) = 0$, the critical cone is given by

$$\begin{aligned} \mathcal{C}_Z^{CC}(x^*, y^*) = \mathcal{L}_Z^{CC}(x^*, y^*) = \{(d_x, d_y) \in \mathbb{R}^n \ : & (d_x)_1 = 0, \\ & (d_y)_1 \leq 0, \quad (d_y)_2 \geq 0, \\ & (d_x)_2 \cdot (d_y)_2 = 0\} \end{aligned}$$

The Hessian in the CC-SOSC condition (3.26) consists only of

$$\nabla^2 f(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

However, we can choose $(d_x, d_y) = ((0,0),(0,1)) \in \mathcal{C}_Z^{CC}(x^*, y^*) \setminus \{0\}$ such that the condition $d_x^T \nabla^2 f(x^*) d_x > 0$ is violated. Thus the (NLP-) Second Order Sufficient Optimality Condition does not hold in this point.

To close this section, we illustrate how the second order sufficiency condition for the reformulation (1.2) can be transferred to the original problem (1.1).

Corollary 3.48. *Let f, g, h be twice continuously differentiable and $x^* \in \mathcal{X}$ be M-stationary.*

(a) If $\|x^*\|_0 = \kappa$ and for all $d_x \in C_{\mathcal{X}}(x^*) \setminus \{0\}$ there exists an M-stationary multiplier (λ, μ, γ) with

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(x^*) \right) d_x > 0.$$

then x^* is a strict local minimum of (1.1).

(b) If $\|x^*\|_0 < \kappa$ but CC-LICQ holds in x^* and the unique M-stationary multiplier (λ, μ, γ) satisfies $\gamma = 0$ as well as

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(x^*) \right) d_x > 0$$

for all $d_x \in C_{\mathcal{X}}(x^*) \setminus \{0\}$, then x^* is a strict local minimum of (1.1).

Proof. (a) If $\|x^*\|_0 = \kappa$, then there exists a unique y^* such that $(x^*, y^*) \in Z$. Also, due to $\|x^*\|_0 = \kappa$ and thus $I_{00}(x^*, y^*) = \emptyset$, M- and S-stationarity coincide and therefore by Theorem 3.46 (x^*, y^*) is a strict local minimum with respect to x of (1.2). And since for all $x \in \mathcal{X} \cap B_r(x^*)$ with $r > 0$ sufficiently small we have $(x, y^*) \in Z \cap B_r(x^*, y^*)$, this implies that x^* is a strict local minimum of (1.1).

(b) Under the given assumptions every point $(x^*, y) \in Z$ is S-stationary with the unique multiplier $(\lambda, \mu, 0)$ and satisfies CC-SOSC. Thus all points (x^*, y) are strict local minima with respect to x of (1.2) on some ball $B_r(x^*, y)$. And since the set $\{(x^*, y) \mid (x^*, y) \in Z\}$ is compact, we can assume without loss of generality that the radius $r > 0$ is independent from y. (A similar argument is used in the proof of Corollary 3.51.) Now assume that x^* was not a strict local minimum of (1.1). Then we could find a sequence $(x^k)_k \subseteq \mathcal{X}$ with $x^k \to x^*$ and $f(x^k) \leq f(x^*)$. For each x^k there exists y^k such that $(x^k, y^k) \in Z$. And for k sufficiently large we have $x^k \in B_r(x^*) \setminus \{x^*\}$ and the implications

$$x_i^* \neq 0 \quad \Longrightarrow \quad x_i^k \neq 0 \quad \Longrightarrow \quad y_i^k = 0.$$

Thus, for all k sufficiently large, we know $(x^*, y^k) \in Z$. Consequently $(x^k, y^k) \in B_r(x^*, y^k)$ and thus $f(x^k) > f(x^*)$.

□

3.3.3 Local Uniqueness of M-stationary Points

While the proofs of Theorems 3.42 and 3.46 cannot be transferred directly to M-stationary points of (1.2), we are able to show that an M-stationary point is locally unique, if CC-CPLD and a second order condition hold. We follow a line of argument by Guo, Lin and Ye [40]. To simplify the presentation of the proof of Theorem 3.50 we show the following auxiliary result first.

Proposition 3.49. *Let $(x^*, y^*) \in Z$ be feasible point of (1.2) and $(x^k, y^k)_k \subseteq Z$ be a sequence of M-stationary points of (1.2) converging to (x^*, y^*).*

(a) If CC-CPLD holds in (x^*, y^*), then (x^*, y^*) is M-stationary and one can find a bounded sequence $(\lambda^k, \mu^k, \gamma^k)_k$ of M-stationary multipliers of (x^k, y^k) such that every accumulation point $(\lambda^*, \mu^*, \gamma^*)$ is an M-stationary multiplier of (x^*, y^*).

(b) If even CC-MFCQ holds in (x^*, y^*), then every sequence $(\lambda^k, \mu^k, \gamma^k)_k$ of M-stationary multipliers of (x^k, y^k) is bounded and every accumulation point $(\lambda^*, \mu^*, \gamma^*)$ is an M-stationary multiplier of (x^*, y^*).

Proof. We begin by verifying (a). Since (x^k, y^k) are M-stationary points of (1.2), there exist multipliers $(\lambda^k, \mu^k, \gamma^k)$ with

$$\nabla f(x^k) + \sum_{i \in I_g(x^k)} \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla h_i(x^k) + \sum_{i \in I_0(x^k)} \gamma_i^k e_i = 0, \tag{3.28}$$

$$\lambda_i^k \geq 0, \quad \lambda_i^k g_i(x^k) = 0, \quad \forall i = 1, \ldots, m, \tag{3.29}$$

$$\gamma_i^k = 0, \quad \forall i \in I_{\pm 0}(x^k, y^k). \tag{3.30}$$

Without loss of generality, we may additionally assume that the vectors

$$\nabla g_i(x^k)\ (i \in \operatorname{supp}(\lambda^k)), \quad \nabla h_i(x^k)\ (i \in \operatorname{supp}(\mu^k)), \quad e_i\ (i \in \operatorname{supp}(\gamma^k)) \tag{3.31}$$

are linearly independent. Otherwise, the multipliers can be modified according to [82, Lemma A.1]. We show that the sequence $(\lambda^k, \mu^k, \gamma^k)_k$ is bounded and thus has a convergent subsequence. To do so, assume for contradiction $\|(\lambda^k, \mu^k, \gamma^k)\| \to \infty$. Then the normed sequence

$$\left(\frac{(\lambda^k, \mu^k, \gamma^k)}{\|(\lambda^k, \mu^k, \gamma^k)\|} \right)_{k \in \mathbb{N}}$$

is bounded and thus (at least on a subsequence) convergent to some nonzero limit $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}) \neq 0$. This limit then satisfies $\bar{\lambda} \geq 0$ and $\bar{\lambda}_i = 0$ for all $i \notin I_g(x^*)$ since then $g_i(x^k) < 0$ and thus $\lambda_i^k = 0$ for all k sufficiently large. Similarly, we know $\bar{\gamma}_i = 0$ for all $i \notin I_0(x^*)$ since then $x_i^k \neq 0$ and thus $\gamma_i^k = 0$ for all k sufficiently large. Additionally, we obtain

$$\sum_{i \in I_g(x^*)} \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^{p} \bar{\mu}_i \nabla h_i(x^*) + \sum_{i \in I_0(x^k)} \bar{\gamma}_i e_i = 0$$

from (3.28). Consequently, the vectors

$$\{\nabla g_i(x), i \in \operatorname{supp}(\bar{\lambda})\} \quad \text{and} \quad \{\nabla h_i(x), i \in \operatorname{supp}(\bar{\mu}),\ e_i i \in \operatorname{supp}(\bar{\gamma})\}$$

are positively linearly dependent in x^* and thus by CC-CPLD have to remain linearly dependent in a neighbourhood. Due to

$$\operatorname{supp}(\bar{\lambda}) \subseteq \operatorname{supp}(\lambda^k), \quad \operatorname{supp}(\bar{\mu}) \subseteq \operatorname{supp}(\mu^k), \quad \operatorname{supp}(\bar{\gamma}) \subseteq \operatorname{supp}(\gamma^k)$$

for all k sufficiently large, we obtain a contradiction to the choice of the multipliers $(\lambda^k, \mu^k, \gamma^k)$ in (3.31).
Thus, the sequence $(\lambda^k, \mu^k, \gamma^k)_k$ is bounded and therefore convergent to some limit $(\lambda^*, \mu^*, \gamma^*)$ on a subsequence.
Since f, g and h are continuously differentiable, we have

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla h_i(x^*) + \sum_{i=1}^{n} \gamma_i^* e_i = 0.$$

Analogously to our previous arguments one sees that $\lambda^* \geq 0$ and $\operatorname{supp}(\lambda^*) \subseteq I_g(x^*)$ as well as $\operatorname{supp}(\gamma^*) \subseteq I_0(x^*)$. Thus, (x^*, y^*) together with the multipliers $(\lambda^*, \mu^*, \gamma^*)$ is M-stationary.
To verify part (b), one only has to observe that, under the assumption of CC-MFCQ, it is not necessary to modify the multipliers to guarantee (3.31) in order to obtain a contradiction. $\square$

The previous result states that the limit of every convergent sequence of M-stationary points is also M-stationary. This plays a major role in the proof of the following uniqueness theorem for M-stationary points. In this result, we need an assumption which is closely related to CC-SOSC, but stronger since condition (3.26) now has to hold for all M-stationary multipliers, not only one S-stationary multiplier. Under CC-LICQ however the M-stationary multiplier is unique. The following result and its proof is motivated by a similar result for MPCCs [40].

Theorem 3.50 (Local uniqueness of M-stationary points)**.** *Let f, g, h be twice continuously differentiable, (x^*, y^*) be an M-stationary point of* (1.2) *satisfying CC-CPLD. Additionally, let*

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i \nabla^2 h_i(x^*) \right) d_x > 0 \tag{3.32}$$

hold for all $(d_x, d_y) \in \mathcal{C}_Z^{CC}(x^*, y^*)$ with $d_x \neq 0$ and <u>*all*</u> *M-stationary multipliers (λ, μ, γ) of (x^*, y^*). Then there exists a radius $r > 0$ such that*

$$\forall (x, y) \in Z \cap B_r(x^*, y^*) : \ [(x, y) \textit{ is M-stationary } \Rightarrow x = x^*] .$$

Proof. Assume that the claim is false. Then there exists a sequence $(x^k, y^k)_{k\in\mathbb{N}} \subset Z$ of M-stationary points converging to (x^*, y^*) with $x^k \neq x^*$. By Proposition 3.49(a) we can assume without loss of generality that the corresponding M-stationary multipliers $(\lambda^k, \mu^k, \gamma^k)$ are convergent, too, and that the limit $(\lambda^*, \mu^*, \gamma^*)$ is an M-stationary multiplier for (x^*, y^*), i.e.

$$\begin{aligned}
&\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla h_i(x^*) + \sum_{i=1}^n \gamma_i^* e_i = 0, \\
&g_i(x^*) \leq 0, \ \lambda_i^* \geq 0, \ \lambda_i^* g_i(x^*) = 0, \quad \forall i = 1, \dots, m, \\
&h_i(x^*) = 0, \quad \forall i = 1, \dots, p, \\
&\gamma_i^* = 0, \quad \forall i \in I_{\pm 0}(x^*, y^*).
\end{aligned} \tag{3.33}$$

For $k \in \mathbb{N}$ define the directions $d^k = (d_x^k, d_y^k)$ by

$$d_x^k := \frac{x^k - x^*}{\|x^k - x^*\|}, \qquad d_y^k := \frac{y^k - y^*}{\|(x^k, y^k) - (x^*, y^*)\|}.$$

We have $\|d_x^k\| = 1$ and $\|d_y^k\| \leq 1$ for all $k \in \mathbb{N}$. Hence the sequences are bounded and we can assume without loss of generality that $d^k = (d_x^k, d_y^k)$ is convergent. Denote the limit by $d = (d_x, d_y)$. Since $\|d_x^k\| = 1$ for all $k \in \mathbb{N}$, we have $d_x \neq 0$.
Furthermore we can show $(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)$ analogously to the proof of Theorem 3.46.
Before we show $\nabla f(x^*)^T d_x \leq 0$, we derive four helpful equations. Since $(\lambda^k, \mu^k, \gamma^k)$ is an M-stationary multiplier for (x^k, y^k), we have

$$\sum_{i=1}^m \lambda_i^k g_i(x^k) + \sum_{i=1}^p \mu_i^k h_i(x^k) + \sum_{i=1}^n \gamma_i^k x_i^k = 0 \tag{3.34}$$

for all $k \in \mathbb{N}$. Because of the continuity of g_i and the properties of the multipliers $(\lambda^k, \mu^k, \gamma^k)$ and $(\lambda^*, \mu^*, \gamma^*)$, the implications

$$\begin{aligned}
g_i(x^*) \neq 0 \ \Rightarrow \ g_i(x^k) \neq 0 \ &\Rightarrow \ \lambda_i^k = 0, \\
x_i^* \neq 0 \ \Rightarrow \ x_i^k \neq 0 \ &\Rightarrow \ \gamma_i^k = 0, \\
\lambda_i^* \neq 0 \ \Rightarrow \ \lambda_i^k \neq 0 \ &\Rightarrow \ g_i(x^k) = 0, \\
\gamma_i^* \neq 0, \ \Rightarrow \ \gamma_i^k \neq 0 \ &\Rightarrow \ x_i^k = 0,
\end{aligned}$$

hold for sufficiently large k. Hence we also have

$$\sum_{i=1}^{m} \lambda_i^k g_i(x^*) + \sum_{i=1}^{p} \mu_i^k h_i(x^*) + \sum_{i=1}^{n} \gamma_i^k x_i^* = 0, \tag{3.35}$$

$$\sum_{i=1}^{m} \lambda_i^* g_i(x^k) + \sum_{i=1}^{p} \mu_i^* h_i(x^k) + \sum_{i=1}^{n} \gamma_i^* x_i^k = 0, \tag{3.36}$$

for all k sufficiently large. Define $\ell : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n \to \mathbb{R}$ by

$$(x, \lambda, \mu, \gamma) \mapsto \ell(x, \lambda, \mu, \gamma) := \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{p} \mu_i h_i(x) + \sum_{i=1}^{n} \gamma_i x_i.$$

A first order Taylor-expansion of $x \mapsto \ell(x, \lambda^k, \mu^k, \gamma^k)$ around x^* evaluated at x^k yields

$$\begin{aligned} 0 &= \ell(x^k, \lambda^k, \mu^k, \gamma^k) \\ &= \ell(x^*, \lambda^k, \mu^k, \gamma^k) + \nabla \ell(x^*, \lambda^k, \mu^k, \gamma^k)^T (x^k - x^*) + \mathrm{o}(\|x^k - x^*\|) \\ &= \nabla \ell(x^*, \lambda^k, \mu^k, \gamma^k)^T (x^k - x^*) + \mathrm{o}(\|x^k - x^*\|) \end{aligned}$$

for sufficiently large k. Here, we used (3.34) and (3.35). By dividing through $\|x^k - x^*\|$ and letting k tend to infinity we get

$$0 = \nabla \ell(x^*, \lambda^*, \mu^*, \gamma^*)^T d_x. \tag{3.37}$$

Using this together with the M-stationarity of (x^*, y^*) we can calculate

$$\begin{aligned} \nabla f(x^*)^T d_x &= -\left(\sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla h_i(x^*) + \sum_{i=1}^{n} \gamma_i^* e_i^T \right)^T d_x \\ &= -\nabla \ell(x^*, \lambda^*, \mu^*, \gamma^*))^T d_x = 0. \end{aligned}$$

Because $(d_x, d_y) \in \mathcal{L}_Z^{CC}(x^*, y^*)$ we consequently have verified $(d_x, d_y) \in \mathcal{C}_Z^{CC}(x^*, y^*)$.
To keep the notation more compact, define ω as an abbreviation for the multipliers $\omega := (\lambda, \mu, \gamma)$. For $k \in \mathbb{N}$ define the functions

$$\begin{aligned} \bar{x}^k : [0,1] \to \mathbb{R}^n, &\quad t \mapsto \bar{x}^k(t) := x^* + t \cdot (x^k - x^*), \\ \bar{\omega}^k : [0,1] \to \mathbb{R}^{m+p+n}, &\quad t \mapsto \bar{\omega}^k(t) := \omega^* + t \cdot (\omega^k - \omega^*), \end{aligned}$$

and $s_k : [0,1] \to \mathbb{R}$ by

$$s_k(t) := \left(\nabla f(\bar{x}^k(t)) + \nabla \ell\left(\bar{x}^k(t), \bar{\omega}^k(t) \right) \right)^T (x^k - x^*) - \ell\left(\bar{x}^k(t), \omega^k - \omega^* \right).$$

Using (3.33)-(3.36) and the fact that $\omega^k = (\lambda^k, \mu^k, \gamma^k)$ is an M-stationary multiplier for

(x^k, y^k) we can calculate

$$\begin{aligned}
s_k(0) &= \left(\nabla f(x^*) + \nabla \ell(x^*, \omega^*)\right)^T (x^k - x^*) - \ell(x^*, \omega^k - \omega^*) \\
&= -\left(\sum_{i=1}^m \lambda_i^k g_i(x^*) + \sum_{i=1}^p \mu_i^k h_i(x^*) + \sum_{i=1}^n \gamma_i^k x_i^*\right) \\
&\quad - \left(\sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{i=1}^p \mu_i^* h_i(x^*) + \sum_{i=1}^p \gamma_i^* x_i^*\right) = 0, \\
s_k(1) &= \left(\nabla f(x^k) + \nabla \ell(x^k, \omega^k)\right)^T (x^k - x^*) - \ell(x^k, \omega^k - \omega^*) \\
&= -\left(\sum_{i=1}^m \lambda_i^k g_i(x^k) + \sum_{i=1}^p \mu_i^k h_i(x^k) + \sum_{i=1}^n \gamma_i^k x_i^k\right) \\
&\quad - \left(\sum_{i=1}^m \lambda_i^* g_i(x^k) + \sum_{i=1}^p \mu_i^* h_i(x^k) + \sum_{i=1}^n \gamma_i^* x_i^k\right) = 0.
\end{aligned}$$

The functions s_k are twice continuously differentiable. The mean value theorem provides the existence of a $\tau_k \in (0, 1)$ such that

$$s_k'(\tau_k) = \frac{s_k(1) - s_k(0)}{1 - 0} = 0. \tag{3.38}$$

Using $(\bar{x}^k)'(t) = x^k - x^*$ and $(\bar{\omega}^k)'(t) = \omega^k - \omega^*$ it is straight forward to calculate

$$s_k'(\tau_k) = (x^k - x^*)^T \left(\nabla^2 f(\bar{x}^k(\tau_k)) + \sum_{i=1}^m \bar{\lambda}_i^k(\tau_k) \nabla^2 g_i(\bar{x}^k(\tau_k)) + \sum_{i=1}^p \mu_i^k(\tau_k) \nabla^2 h_i(\bar{x}^k(\tau_k))\right) (x^k - x^*).$$

Since τ_k is bounded we have $\bar{x}^k(\tau_k) \to x^*$, $\bar{\omega}^k(\tau_k) \to \omega^*$ for $k \to \infty$. It follows from (3.38) that

$$\frac{(x^k - x^*)^T}{\|x^k - x^*\|} \left(\nabla^2 f(\bar{x}^k(\tau_k)) + \sum_{i=1}^m \bar{\lambda}_i^k(\tau_k) \nabla^2 g_i(\bar{x}^k(\tau_k)) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(\bar{x}^k(\tau_k))\right) \frac{(x^k - x^*)}{\|x^k - x^*\|} = 0$$

for sufficiently large $k \in \mathbb{N}$ and thus for $k \to \infty$

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(x^*)\right) d_x = 0,$$

since the functions f, g and h are twice continuously differentiable. Because $(d_x, d_y) \in \mathcal{C}_Z^{CC}(x^*, y^*)$ and $d_x \neq 0$, this is a contradiction to the theorem's assumption. $\square$

The condition in Theorem 3.50 is different from the CC-SOSC, confirm Definition 3.44. Instead of requiring (3.26) to hold for one triple of multipliers, Theorem 3.50 requires the inequality (3.32) to hold for all multipliers of the given M-stationary point.
Since the definition of an M-stationary point is independent from y, we can formulate a result on uniqueness of M-stationary points directly for (1.1).

Corollary 3.51. *Let f, g, h be twice continuously differentiable. Let x^* be feasible for* (1.1)*, M-stationary, satisfy CC-CPLD and let*

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x^*) \right) d_x > 0$$

hold for all $d_x \in \mathcal{C}_{\mathcal{X}}(x^*)$ with $d_x \neq 0$ and <u>all</u> M-stationary multipliers (λ, μ, γ) of x^*. Then there exists a radius $r > 0$ such that

$$\forall x \in \mathcal{X} \cap B_r(x^*) : \ [x \textit{ is M-stationary } \Rightarrow x = x^*] \, .$$

Proof. For every $\bar{y}$, such that $(x^*, \bar{y}) \in Z$, the point $(x^*, \bar{y})$ is M-stationary for (1.2). Due to the definition of $d_x \in \mathcal{C}_{\mathcal{X}}(x^*)$ the prerequisites of Theorem 3.50 are satisfied and thus there exists $r_{\bar{y}} > 0$ such that

$$\forall (x, y) \in Z \cap B_{r_{\bar{y}}}(x^*, \bar{y}) : \ [(x, y) \text{ is M-stationary } \Rightarrow x = x^*] \, .$$

Together the balls $B_{r_y}(x^*, y)$ form an open covering of the compact set $\{(x, y) \in Z \mid x = x^*\}$ and thus we can find an $r > 0$ such that

$$\forall (x, y) \in Z \cap B_r(x^*, \bar{y}) : \ [(x, y) \text{ is M-stationary } \Rightarrow x = x^*]$$

for all $(x^*, \bar{y}) \in Z$.
Now consider an arbitrary M-stationary point $x \in \mathcal{X} \cap B_r(x^*)$. Then for all y such that $(x, y) \in Z$ the point (x, y) is M-stationary for (1.2). By choosing $r > 0$ sufficiently small we can ensure the implication

$$x_i^* \neq 0 \quad \Longrightarrow \quad x_i \neq 0 \quad \Longrightarrow \quad y_i = 0$$

and thus obtain $(x^*, y) \in Z$. This implies $(x, y) \in B_r(x^*, y)$ and thus $x = x^*$. □

We apply the above result in Section 4.3.1 to ensure the local convergence of a Scholtes-type regularisation method.

3.4 Optimality Conditions for the Cardinality Constrained Problem

In this section we compare the optimality conditions from the previous sections to related results. Some of the results presented in this chapter, for example Corollary 3.43, can be seen as results for the cardinality constrained problem (1.1) directly. Some recent results for optimality conditions for (1.1) are derived directly from that problem and do not use auxiliary variables. We discuss two approaches that use techniques from nonlinear optimization [6, 71] and investigate how they relate to the concepts for (1.2).
In [6] the cardinality constrained optimization problem is considered without further constraints, i.e.

$$\min_{x \in \mathbb{R}^n} \ f(x) \quad \text{s.t.} \quad \|x\|_0 \leq \kappa. \tag{3.39}$$

Three different optimality conditions are introduced and discussed. Problem (3.39) is the cardinality constrained problem (1.1) in case there are no additionally equality or inequality constraints. Thus we have

$$\mathcal{X} = \{x \in X \ : \ \|x\|_0 \leq \kappa\} = \{x \in \mathbb{R}^n \ : \ \|x\|_0 \leq \kappa\}.$$

In [6] Beck and Eldar define the following three conditions, which we compare to S- and M-stationarity in the following discussion.

Definition 3.52 ([6])**.** Let x^* be feasible for (3.39).

(a) We call x^* a *basic feasible (BF) vector*

(I) if $\nabla f(x^*) = 0$, in case $\|x^*\|_0 < \kappa$,

(II) if $(\nabla f(x^*))_j = 0$ for all $j \in \operatorname{supp}(x^*)$, in case $\|x^*\|_0 = \kappa$.

(b) If there exists an $L > 0$ such that

$$x^* \in P_{\mathcal{X}}\left(x^* - \frac{1}{L}\nabla f(x^*)\right),$$

we call x^* an *L-stationary point.*

(c) We call x^* a *CW-minimum*

(I) in case $\|x^*\|_0 < \kappa$: If for all $i = 1, \ldots, n$: $f(x^*) = \min_{t\in\mathbb{R}} f(x^* + t \cdot e_i)$,

(II) in case $\|x^*\|_0 = \kappa$: If for all $j \in \operatorname{supp}(x^*)$ and for all $i = 1, \ldots, n$:

$$f(x^*) \leq \min_{t\in\mathbb{R}} f(x^* - x_j^* \cdot e_j + t \cdot e_i).$$

The following implications between the conditions for a BF vector, an L-stationary point and a CW-minimum hold.

Proposition 3.53 ([6, Corollary 2.1, Theorem 2.4, Lemma 2.5])**.** *Let x^* be feasible for* (3.39)*. Then:*

1. *If x^* is L-stationary for an $L > 0$, then x^* is a BF vector.*

2. *If x^* is a CW-minimum, then x^* is a BF-vector.*

3. *If x^* is a CW-minimum and ∇f is Lipschitz-continuous with Lipschitz-constant $L(f)$, then x^* is $L(f)$-stationary.*

The conditions for a BF vector, L-stationary and a CW-minimum are necessary optimality conditions for (3.39), as stated by the following theorem.

Theorem 3.54 ([6, Theorems 2.1, 2.2, 2.3])**.** *Let x^* be a solution of* (3.39)*. Then:*

(a) The point x^* is a CW-minimum (and therefore, by Proposition 3.53, also a BF vector)

(b) If additionally ∇f is Lipschitz-continuous with Lipschitz-constant $L(f)$, then x^* is L-stationary for every $L > L(f)$ and

$$P_{\mathcal{X}}\left(x^* - \frac{1}{L}\nabla f(x^*)\right)$$

is single valued.

We will now discuss the relations between M- and S-stationarity and the concepts from [6]. The complementarity formulation of (3.39) is given by

$$\begin{aligned} \min_{x,y} \; f(x) \quad \text{s.t.} \quad & e^T y \geq n - \kappa, \\ & 0 \leq y \leq 1, \\ & x \circ y = 0. \end{aligned} \tag{3.40}$$

The following relations between BF vectors and M- and S-stationary points hold.

Theorem 3.55. *Let x^* be feasible for* (3.39) *and $y^* \in \mathbb{R}^n$ such that (x^*, y^*) is feasible for* (3.40) *Then:*

(a) If (x^*, y^*) is M-stationary and $\|x^*\|_0 = \kappa$, then x^* is a BF vector.

(b) If x^* is a BF vector, then for every $y \in \mathbb{R}^n$ such that (x^*, y) is feasible for (3.40)*, the point (x^*, y) is S-stationary.*

Proof. (a) Let (x^*, y^*) be feasible for (3.40) and M-stationary. Then $\nabla f(x^*) \in \text{span}\{e_i : i \in I_0(x^*)\}$. Consequently $(\nabla f(x^*))_i = 0$ for all $i \in \text{supp}(x^*)$. Since $\|x^*\|_0 = \kappa$, this means that x^* is a BF vector.

(b) Let x^* be a BF vector.

In case $\|x^*\|_0 < \kappa$ holds, we have $\nabla f(x^*) = 0$. Thus for any $y \in \mathbb{R}^n$, such that (x^*, y) is feasible for (3.40), we have that (x^*, y) is S-stationary with multipliers $\gamma = 0$.

In case $\|x^*\|_0 = \kappa$, we have $I_{00}(x^*, y) \cup I_{0+}(x^*, y) = \emptyset$. Since x^* is a BF-vector, we have $\nabla f(x^*) \in \text{span}\{e_i : i \in I_0(x^*)\}$, thus in fact $\nabla f(x^*) \in \text{span}\{e_i : I_{01}(x^*, y)\}$. This means we have $\nabla f(x^*) + \sum_{i=1}^n \gamma_i e_i = 0$ for some γ_i, $i \in I_{01}(x^*, y^*)$, hence any point $(x^*, y) \in Z$ is S-stationary.

□

By Proposition 3.53 a CW-minimum is a BF vector. Hence by the above Theorem a feasible CW-minimum x^* is also S-stationary, in the sense that every point (x^*, y) feasible for (3.40) is S-stationary. For the above implication of the BF vector property by M-stationarity, it is necessary that the cardinality constraint is active. Furthermore, S-stationarity of a point (x^*, y^*) does not imply that x^* is a CW-minimum (even if the cardinality constraint is active). For these statements we give a short example.

Example 3.56. Let $n = 3$, $\kappa = 1$ and consider $f(x) = (x_1 - 1)^2 + x_3$. Then $\nabla f(x) = (2 \cdot (x_1 - 1), 0, 1)^T$.
Let $x^* = (0, 0, 0)$ and $y^* = (1, 1, 1)$. Then (x^*, y^*) is feasible for (3.40) and $I_{01}(x^*, y^*) = \{1, 2, 3\}$. We have $\nabla f(x^*) = (-2, 0, 1)^T$. Choosing $\gamma^* = (2, 0, -1)$ we have

$$\nabla f(x^*) + \sum_{i \in I_{0+1} \cup I_{01}} \gamma_i^* e_i = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} + \sum_{i \in I_{01}} \gamma_i^* e_i = 0.$$

Hence (x^*, y^*) is S-stationary. However, since $\nabla f(x^*) \neq 0$ and $\|x^*\|_0 < \kappa$, the point x^* is not a BF vector.

Now consider the point $\hat{x} = (1,0,0)$ and let $\hat{y} = (0,1,1)$. Then $(\hat{x},\hat{y})$ is feasible for (3.40) and $I_{01} = \{2,3\}$. We have $\nabla f(\hat{x}) = e_3$ and thus choosing the multiplier $\hat{\gamma} = (0,0,-1)$ the point $(\hat{x},\hat{y})$ is S-stationary. We have $\|\hat{x}\|_0 = \kappa$ and $f(\hat{x}) = 0$. Let $j = 1 \in \operatorname{supp}(\hat{x})$, $i = 3$ and $t \in \mathbb{R}$. Then

$$f(\hat{x} - \hat{x}^*_j e_j + t \cdot e_i) = f(0,0,t) = 1 + t \to -\infty < 0 = f(\hat{x})$$

as $t \to -\infty$. Thus the point $\hat{x}$ is not a CW-minimum.

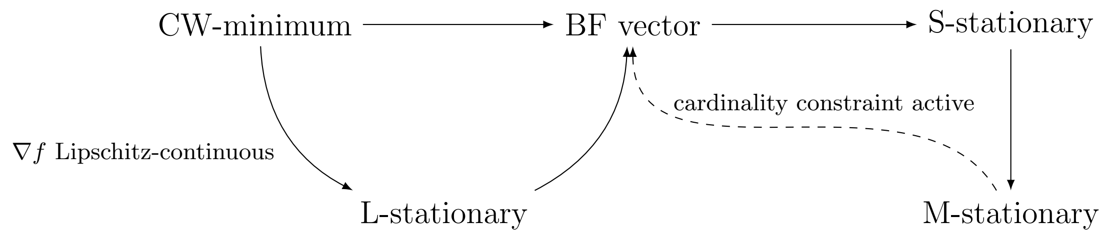


Figure 3.3: Relations between stationary conditions for (3.39).

The above example concludes our comparison of CW minima, L-stationary points and BF vectors with M- and S-stationary points for the special case (3.39). The stationarity conditions from [6] are stronger than M- and S-stationarity for the complementarity formulation, for the case $X = \mathbb{R}^n$. However, the constraints of the set X play a central role in the analysis of (1.2) and the derivation of M- and S-stationarity. Since the conditions for a BF vector imply S- and therefore M-stationarity, we can conclude that the two approaches are consistent. The relations are illustrated in Figure 3.3.

Another nonlinear programming approach to derive optimality conditions for (1.1) was given in [71]. The authors consider the cardinality constrained optimization problem directly. First order optimality conditions are derived using the Fréchet-, Limiting- and Clarke-normal cones of the set

$$\mathcal{X}_\kappa := \{x \in \mathbb{R}^n \,:\, \|x^*\|_0 \leq \kappa\}.$$

We repeat the definition of these cones.

Definition 3.57 (Fréchet normal cone [77] and limiting normal cone [65, Definition 1.1])**.** Let $A \subseteq \mathbb{R}^n$ be nonempty and $x \in A$.

(a) Let A be closed. The set

$$\mathcal{N}^M_A(x) := \{w \in \mathbb{R}^n \,:\, \exists (x^k)_{k\in\mathbb{N}}, (w^k)_{k\in\mathbb{N}} : \; x^k \to x,\; w^k \to w,\; w^k \in \mathcal{N}^F_A(x^k)\ \forall k \in \mathbb{N}\}$$

is called *limiting normal cone of Z at x*. For $x \notin A$ set $\mathcal{N}^M_A(x) := \emptyset$.

(b) The *Clarke normal cone to A at x^** is defined as $\mathcal{N}^C_A(x^*) := \operatorname{conv}(\mathcal{N}^M_A(x^*))$.

(c) The set

$$\mathcal{N}^F_A(x) := \mathcal{T}_A(x)^\circ$$

is called *Fréchet normal cone of Z at x*. For $x \notin A$ set $\mathcal{N}^F_A(x) := \emptyset$.

Using the Clarke normal cone as well as the Fréchet- and Limiting-normal cones, the following conditions were introduced in [71].

Definition 3.58. Let x^* be feasible for (1.1). If there exist $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ such that

$$-\nabla f(x^*) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) - \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) \in \mathcal{N}^F_{\mathcal{X}_\kappa}(x^*) \text{ (or } \mathcal{N}^M_{\mathcal{X}_\kappa}(x^*), \text{ or } \mathcal{N}^C_{\mathcal{X}_\kappa}(x^*)),$$
$$\lambda_i \geq 0, \quad \lambda_i \cdot g_i(x^*) = 0 \quad \forall i = 1, \dots, m,$$

then x^* is called a *B-KKT* (or *M-KKT*, or *C-KKT*) point of (1.1).

The following lemma states representations of the normal cones that are used to define the above stationarity conditions. Keeping these in mind, it is easy to see that B-KKT points are also M-KKT points, which are in turn also C-KKT points. Furthermore, if there are no equality or inequality constraints present, this means we have a cardinality constrained problem of the form (3.39), B-KKT points and BF vectors coincide.

Lemma 3.59 ([71, Lemma 2.3], cf. also [4, Theorem 3.9, Theorem 3.15] and [72, Theorem 2.1, Theorem 2.2])**.** *Let $x^* \in \mathcal{X}_\kappa$. We have*

$$\mathcal{T}_{\mathcal{X}_\kappa}(x^*) = \begin{cases} \operatorname{span}\{e_i \,:\, i \in \operatorname{supp}(x^*)\}, & \textit{if } \|x^*\|_0 = \kappa, \\ \bigcup_{J \in \mathcal{J}} \operatorname{span}\{e_i \,:\, i \in J\}, & \textit{if } \|x^*\|_0 < \kappa, \end{cases}$$
$$\mathcal{N}^F_{\mathcal{X}_\kappa}(x^*) = \begin{cases} \operatorname{span}\{e_i \,:\, i \in I_0(x^*)\}, & \textit{if } \|x^*\|_0 = \kappa, \\ \{0\}, & \textit{if } \|x^*\|_0 < \kappa, \end{cases}$$
$$\mathcal{N}^M_{\mathcal{X}_\kappa}(x^*) = \begin{cases} \operatorname{span}\{e_i \,:\, i \in I_0(x^*)\}, & \textit{if } \|x^*\|_0 = \kappa, \\ \bigcup_{J \in \mathcal{J}} \operatorname{span}\{e_i \,:\, i \in J^C\}, & \textit{if } \|x^*\|_0 < \kappa, \end{cases}$$
$$\mathcal{N}^C_{\mathcal{X}_\kappa}(x^*) = \operatorname{span}\{e_i \,:\, i \in I_0(x^*)\},$$

where $\mathcal{J} := \{J \subseteq \{1, \dots, n\} \,:\, \operatorname{supp}(x^*) \subseteq J, \ |J| = \kappa\}$.

Following this approach, LICQ- and MFCQ-type constraint qualifications for (1.1) can be defined.

Definition 3.60. Let x^* be feasible for (1.1). We say that x^* satisfies

(a) *R-LICQ*, if

- in case $\|x^*\|_0 = \kappa$: The vectors $\nabla g_i(x^*)$, $i \in I_g(x^*)$, $\nabla h_i(x^*)$, $i = 1, \dots, p$ are linearly independent,
- in case $\|x^*\|_0 < \kappa$: The vectors $(\nabla g_i(x^*))_{\operatorname{supp}(x^*)}$, $i \in I_g(x^*)$, $(\nabla h_i(x^*))_{\operatorname{supp}(x^*)}$, $i = 1, \dots, p$ are linearly independent.

(b) *R-MFCQ*, if

- in case $\|x^*\|_0 = \kappa$: The vectors $\nabla h_i(x^*)$, $i = 1, \dots, p$ are linearly independent, and there exists a $d \in \mathbb{R}^n$ such that

$$\nabla g_i(x^*)^T d < 0, \ \forall i \in I_g(x^*), \ \nabla h_i(x^*)^T d = 0, \ \forall i = 1, \dots, p.$$

- in case $\|x^*\|_0 < \kappa$: The vectors $(\nabla h_i(x^*))_{\text{supp}(x^*)}$, $i = 1, \ldots, p$ are linearly independent, and for each $J \in \{\hat{J} \subseteq \{1, \ldots, n\} : \text{supp}(x^*) \subseteq \hat{J}, |\hat{J}| = \kappa\}$ there exists a $d \in \text{span}\{e_i : i \in J\}$ such that

$$\nabla g_i(x^*)^T d < 0, \; \forall i \in I_g(x^*), \; \nabla h_i(x^*)^T d = 0, \; \forall i = 1, \ldots, p.$$

Under these constraint qualifications a local minimum of (1.1) is an M-KKT point or even a B-KKT point.

Theorem 3.61 ([71, Theorem 3.2, Theorem 3.4])**.** *Let x^* be a local minimum of* (1.1)*.*

(a) If R-LICQ holds at x^*, then x^* is a B-KKT point.

(b) If R-MFCQ holds at x^*, then x^* is an M-KKT point.

We will now discuss how the above concepts relate to the constraint qualifications and stationary conditions for (1.2).
The constraint qualifications R-LICQ and R-MFCQ are equivalent to CC-LICQ and CC-MFCQ, as the following proposition states. By distinguishing the two cases of an active and inactive cardinality constraint and using Motzkin's theorem of the alternative (see [60]), it is easy to verify the following result. Because it is straightforward, we omit the proof.

Proposition 3.62. *Let x^* be feasible for* (1.1)*. Then:*

(a) CC-LICQ holds at x^* if and only if R-LICQ holds at x^*,

(b) CC-MFCQ holds at x^* if and only if R-MFCQ holds at x^*.

The conditions for B-, M- and C-KKT points of (1.1) and S- and M-stationary points of (1.2) can also be put into relation.

Proposition 3.63. *Let x^* be feasible for* (1.1)*.*

(a) If (x^*, y^*) is an M-stationary point of (1.2) *and $\|x^*\|_0 = \kappa$, then x^* is a B-KKT point of* (1.1)*.*

(b) Let x^* be an M-KKT point of (1.1)*. Then there exists a vector $y \in \mathbb{R}^n$ such that (x^*, y) is an S-stationary point of* (1.2)*.*

(c) The point x^* is a C-KKT point of (1.1) *if and only if x^* is an M-stationary point of* (1.2)*.*

Proof. (a) Let (x^*, y^*) be an M-stationary point with multipliers (λ, μ, γ) and assume $\|x^*\|_0 = \kappa$. Then $-\nabla f(x^*) - \sum_{i \in I_g(x^*)} - \sum_{i=1}^p \mu_i \nabla h_i(x^*) \in \text{span}\{e_i : i \in I_0(x^*)\} = \mathcal{N}^F_{\mathcal{X}_\kappa}(x^*)$, hence x^* is a B-KKT point.

(b) Let x^* be an M-KKT point of (1.1). Then there are multipliers $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p$ such that $\lambda \geq 0$, $\lambda_i \cdot g_i(x^*) = 0$ for all $i \in I_g(x^*)$ and

$$-\nabla f(x^*) - \sum_{i \in I_g} \lambda_i \nabla g_i(x^*) - \sum_{i=1}^p \mu_i \nabla h_i(x^*) \in \mathcal{N}^M_{\mathcal{X}_\kappa}(x^*)$$

$$= \begin{cases} \text{span}\{e_i : i \in I_0(x^*)\}, & \text{if } \|x^*\|_0 = \kappa, \\ \bigcup\limits_{J \in \mathcal{J}} \text{span}\{e_i : i \in J^C\}, & \text{if } \|x^*\|_0 < \kappa, \end{cases}$$

where $\mathcal{J} = \{J \subseteq \{1, \dots, n\} : \operatorname{supp}(x^*) \subseteq J,\ |J| = \kappa\}$. In case $\|x^*\|_0 = \kappa$, let

$$y_i^* := \begin{cases} 0, & \text{if } x_i^* \neq 0, \\ 1, & \text{if } x_i^* = 0, \end{cases} \tag{3.41}$$

for $i = 1, \dots, n$. We have $(x^*, y^*) \in Z$, $I_{00}(x^*, y^*) = \emptyset$ and thus

$$\begin{aligned} -\nabla f(x^*) - \sum_{i \in I_g} \lambda_i \nabla g_i(x^*) - \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) &\in \operatorname{span}\{e_i : i \in I_0(x^*)\} \\ &= \operatorname{span}\{e_i : i \in I_{0+}(x^*, y^*) \cup \in I_{01}(x^*, y^*)\}. \end{aligned}$$

Hence (x^*, y^*) is S-stationary.

In case $\|x^*\|_0 < \kappa$, we have

$$-\nabla f(x^*) - \sum_{i \in I_g(x^*)} \lambda_i \nabla g_i(x^*) - \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) \in \bigcup_{J \in \mathcal{J}} \operatorname{span}\{e_i : i \in J^C\}.$$

Hence there is $\tilde{J} \subseteq I_0(x^*)$ and coefficients γ_i, $i \in \tilde{J}$, such that $|\tilde{J}| = n - \kappa$ and

$$-\nabla f(x^*) - \sum_{i \in I_g(x^*)} \lambda_i \nabla g_i(x^*) - \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) = \sum_{i \in \tilde{J}} \gamma_i e_i.$$

Choosing y^* according to (3.41), we have $(x^*, y^*) \in Z$ and $I_0(x^*) = I_{01}(x^*, y^*)$. Since $\tilde{J} \subseteq I_0(x^*)$, the point (x^*, y^*) is S-stationary.

(c) Let $\lambda \in \mathbb{R}^m$, $\lambda \geq 0$, $\mu \in \mathbb{R}^p$ and $\gamma \in \mathbb{R}^n$. Because by Lemma 3.59 we have $\mathcal{N}^C_{\mathcal{X}_\kappa}(x^*) = \operatorname{span}\{e_i : i \in I_0(x^*)\}$, the condition

$$-\nabla f(x^*) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) - \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) \in \mathcal{N}^C_{\mathcal{X}_\kappa}(x^*)$$

is equivalent to

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i \in I_0(x^*)} \gamma_i e_i = 0,$$

for some coefficients γ_i. Hence the conditions for x^* being a C-KKT point and for being an M-stationary point coincide.

□

For an illustration of the above result see Figure 3.4. If the cardinality constraint is active, the implication indicated by the dashed arrow holds. In this case all conditions coincide.
For the x-part of an S-stationary point to be a B-KKT point, it is in general necessary that the cardinality constraint is active: The S-stationary point (x^*, y^*) in Example 3.56 is not a B-KKT point. The following example shows that in this case the x-part cannot be expected to be an M-KKT point either.

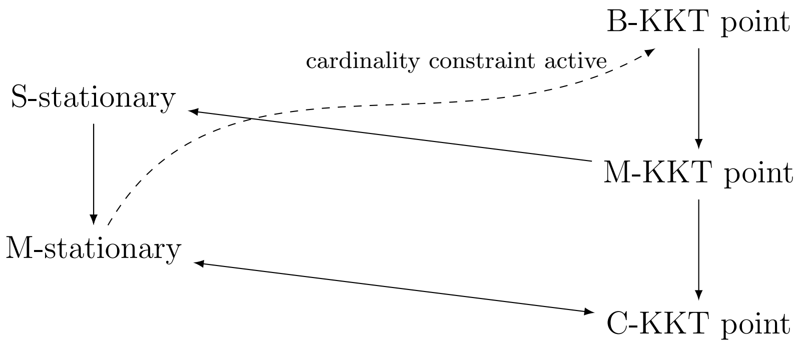


Figure 3.4: Relations between B-, M- and C-KKT points of (1.1) with S- and M-stationary points of (1.2).

Example 3.64. Let $f(x) := (x_1 - 1)^2 + (x_2 - 1)^2$, $x \in \mathbb{R}^2$ and consider the complementarity formulation

$$\min_{x \in \mathbb{R}^2, y \in \mathbb{R}^2} f(x) \quad \text{s.t.} \quad 0 \leq y \leq e,\ e^T y \geq 1,\ x \circ y = 0.$$

Let $x^* = (0, 0)$ and $y^* = (\frac{1}{2}, \frac{1}{2})$. We then have $\|x^*\|_0 = 0 < 1$ and $\nabla f(x^*) = (-2, -2)^T$. The point (x^*, y^*) is an S-stationary point with multipliers $\gamma_1 := \gamma_2 := 2$. Yet we have

$$\nabla f(x^*) = \begin{pmatrix} -2 \\ -2 \end{pmatrix} \notin \operatorname{span}\{e_1\} \cup \operatorname{span}\{e_2\} = \mathcal{N}^M_{\mathcal{X}_\kappa}(x^*).$$

Thus x^* is not an M-KKT point (and therefore also not a B-KKT point).

In addition, second order optimality conditions for (1.1) are considered in [71]. The sufficient optimality condition [71, Theorem 4.2] is analogous to Theorem 3.13: Under a second order condition a B-KKT point is a strict local minimum of (1.1). The CC-SOSC presented in Section 3.3.2 on the other hand takes the special structure of the complementarity-formulation into account: If the CC-SOSC is satisfied in a given S-stationary point of (1.2), we have seen that this point is a strict local minimum with respect to x. Thus these results are related, but not directly comparable.

Let us consider the second order necessary optimality condition given in [71, Theorem 4.1] (Theorem 3.65 below). For the remainder of this section we assume f, g and h to be twice continuously differentiable. To formulate the second order necessary condition we will use a subset of X. Given a B-KKT point x^* with multipliers (λ^*, μ^*), let

$$\begin{aligned} \tilde{X} :=& \{x \in X \ : \ g_i(x) = 0 \ \forall i : \ \lambda_i^* > 0\} \\ =& \{x \in \mathbb{R}^n \ : \ g_i(x) = 0 \ \forall i : \ \lambda_i^* > 0; \ g_i(x) \leq 0 \ \forall i : \ \lambda_i^* = 0; \ h_i(x) = 0 \ \forall i = 1, \ldots, p\}. \end{aligned} \tag{3.42}$$

Note that the set $\tilde{X}$ depends on the B-KKT multipliers (λ^*, μ^*) of x^*. We repeat the second order necessary optimality condition for the cardinality constrained problem (1.1).

Theorem 3.65 ([71, Theorem 4.1])**.** *Let f, g and h be twice continuously differentiable. Let x^* be a local minimum of* (1.1) *and let R-LICQ hold at x^*. Then x^* is a B-KKT point. Let*

(λ^*, μ^*) *be its multipliers. We have*

$$d^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla h_i(x^*) \right) d \geq 0 \tag{3.43}$$

for all $d \in \mathcal{T}_{\tilde{X}}(x^*) \cap \mathcal{T}_{\mathcal{X}_\kappa}(x^*)$.

We now compare condition (3.43) in Theorem 3.65 to Corollary 3.43. To this end let x^* be a local minimum of (1.1) and let CC-LICQ hold at x^*. Then x^* is an S-stationary point and by Proposition 3.34 all S-stationary points (x^*, y) share a unique multiplier $(\bar{\lambda}, \bar{\mu}, \bar{\gamma})$. Note that CC-LICQ is equivalent to R-LICQ. Since x^* is a local minimum, we have that x^* is also a B-KKT point. It is easy to check that the corresponding multipliers coincide due to CC-LICQ. Thus in fact we have $\bar{\lambda} = \lambda^*$ and $\bar{\mu} = \mu^*$.
By Corollary 3.43 we have

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(x^*) \right) d_x \geq 0 \tag{3.44}$$

for all $d_x \in \mathcal{C}_{\mathcal{X}}(x^*)$. Using (3.23) and the fact that x^* is S-stationary, we can calculate a representation of $\mathcal{C}_{\mathcal{X}}(x^*)$ that uses the multipliers $(\lambda^*, \mu^*, \gamma^*)$:

$$\begin{aligned}
\mathcal{C}_{\mathcal{X}}(x^*) = \{d_x \in \mathcal{L}_{\mathcal{X}}(x^*) \, : \quad & \nabla g_i(x^*)^T d_x = 0 \ \forall I_{g+}(x^*, \lambda^*)\} \\
= \{d_x \in \mathbb{R}^n \, : \quad & \nabla g_i(x^*)^T d_x = 0 \ \forall I_{g+}(x^*, \lambda^*), \\
& \nabla g_i(x^*)^T d_x \leq 0 \ \forall i \in I_{g0}(x^*, \lambda^*), \\
& \nabla h_i(x^*)^T d_x = 0 \ \forall i = 1, \dots, p, \\
& |\{i \in I_0(x^*) \, : \, (d_x)_i = 0\}| \geq n - \kappa\} \\
= \{d_x \in \mathcal{L}_{\tilde{X}}^{NLP}(x^*) \, : \quad & |\{i \in I_0(x^*) \, : \, (d_x)_i = 0\}| \geq n - \kappa\}.
\end{aligned}$$

The first equality can be shown exactly like in the proof of Proposition 3.41. The second equality holds, because (due to CC-LICQ) the B-KKT multipliers of x^* in (3.42) coincide with its S-stationary multipliers, which we use in the above representation of $\mathcal{C}_{\mathcal{X}}(x^*)$.
Now let $d_x \in \mathcal{T}_{\tilde{X}}(x^*) \cap \mathcal{T}_{\mathcal{X}_\kappa}(x^*)$ be arbitrary. We distinguish two cases. Firstly, let $\|x^*\|_0 = \kappa$. Then $\mathcal{T}_{\mathcal{X}_\kappa}(x^*) = \text{span}\{e_i \, : \, i \in \text{supp}(x^*)\}$ and therefore $(d_x)_i = 0$ for all $i \in I_0(x^*)$. We thus have

$$|\{i \in I_0(x^*) \, : \, (d_x)_i = 0\}| = |I_0(x^*)| \geq n - \kappa.$$

In fact, equality holds in the last inequality. Secondly, let $\|x^*\|_0 < \kappa$. Then

$$\mathcal{T}_{\mathcal{X}_\kappa}(x^*) = \bigcup_{J \in \mathcal{J}} \text{span}\{e_i \, : \, i \in J\},$$

where $\mathcal{J} = \{J \subseteq \{1, \dots, n\} \, : \, J \supseteq \text{supp}(x^*), \, |J| = \kappa\}$. Hence there exists $\hat{J} \subseteq \{1, \dots, n\}$ such that $\text{supp}(x^*) \subset \hat{J}$, $|\hat{J}| = \kappa$ and $d_x \in \text{span}\{e_i \, : \, i \in \hat{J}\}$. Consequently $\hat{J}^C \subseteq I_0(x^*)$ and $(d_x)_i = 0$ for all $i \in \hat{J}^C$, and thus

$$|\{i \in I_0(x^*) \, : \, (d_x)_i = 0\}| \geq |\{i \in \hat{J}^C \, : \, (d_x)_i = 0\}| = |\hat{J}^C| = n - \kappa.$$

Additionally, in both cases, we have $\mathcal{T}_{\tilde{X}}(x^*) \subseteq \mathcal{L}^{NLP}_{\tilde{X}}(x^*)$ and therefore $d_x \in \mathcal{C}_{\mathcal{X}}(x^*)$. Because by Corollary 3.43 we have (3.44), condition (3.43) is satisfied.

In this section we compared optimality conditions for the complementarity formulation (1.2) with optimality conditions formulated for the cardinality constrained optimization problem (1.1) directly. The conditions given in [6] are only applicable in case there are no further constraints besides the cardinality constraint. We have seen that they are consistent in this case: The conditions from [6] imply S- and M-stationarity, as formalised in Theorem 3.54. The optimality conditions given in [71] are applicable to (1.1) and correspond to some extent to the optimality conditions given in Section 3.2 and Section 3.3. The (discontinuous) cardinality constraint remains as a constraint in the conditions from [71]. Yet both approaches lead to results which can be put in relation to each other. The constraint qualifications CC-LICQ and CC-MFCQ are equivalent to R-LICQ and R-MFCQ. The different stationary conditions coincide, if the cardinality constraint is active. This is not the case, if the cardinality constraint is inactive, as shown by Example 3.64. Additionally the conditions for S-stationarity and M-stationarity hold under weaker constraint qualifications: By Theorem 3.32, for a local minimum of (1.2) to be an S-stationary point, the CC-CPLD constraint qualification is required to hold. For a local minimum of (1.1) to be an M- or C-KKT point the stronger R-MFCQ is required. Moreover, the benefit of an analysis of the continuous reformulation are optimality conditions which are numerically exploitable with nonlinear programming methods [11, 13, 14, 26]. Regarding second order necessary optimality conditions, the two approaches lead to very similar results: Theorem 3.65 is captured by Corollary 3.43.

3.5 A Local Error Bound for the Complementarity Formulation

In this section we will derive a local error bound for the feasible set of (1.2). This result is from the preprint [12], which is in preparation. Given an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^n$, the distance between a set $A \subseteq \mathbb{R}^n$ and a point $x \in \mathbb{R}^n$ is defined as

$$\operatorname{dist}_A(x) := \inf_{z \in A} \|x - z\|.$$

An error bound is an upper and a lower bound on this distance. Given two sets $A, B \subseteq \mathbb{R}^n$, an error bound on the pair (A, B) is given by inequalities

$$c_1 \cdot r(x)^{\gamma_1} \leq \operatorname{dist}_x(A) \leq c_2 \cdot r(x)^{\gamma_2} \quad \forall x \in B,$$

where $c_1, c_2, \gamma_1, \gamma_2 > 0$ and $r : A \cup B \to [0, \infty)$ is a given *residual function* that satisfies

$$r(x) = 0 \quad \Leftrightarrow \quad x \in A,$$

compare [73]. In case $B = \mathbb{R}^n$ the error bound is said to be *global*. Applications of error bounds in mathematical programming include the derivation of rates of convergence and stopping rules for iterative algorithms, see [73]. In particular we will apply Theorem 3.69, the main result of this section, to derive an exact penalty function in Chapter 4. We will use the following definition of a *local* error bound.

Definition 3.66. Let $M \subseteq \mathbb{R}^n$ and $\bar{x} \in M$. We say M *has a local error bound at* $\bar{x}$, if there exists a constant $c > 0$, a neighbourhood $N(\bar{x})$ of $\bar{x}$ and a function $r : \mathbb{R}^n \to [0, \infty)$ such that

$$\operatorname{dist}_M(x) \leq c \cdot r(x) \quad \forall x \in N(\bar{x}),$$

and $r(x) = 0$ if and only if $x \in M$.

To derive the subsequent result, we will use the piecewise decomposition from Section 3.2.1 to prove that a local error bound holds under CC-CPLD. The line of argument is inspired by similar results for Mathematical Programs with Equilibrium Constraints [18] and Generalised Nash Equilibrium Problems [49]. We begin with a relation between error bounds of the piecewise problems and an error bound for the relaxation (1.2).

Lemma 3.67. *If NLP(J) has a local error bound at* (x^*, y^*) *for every* $J \subseteq I_{00}(x^*, y^*)$ *with*

$$\begin{aligned} r_J(x,y) := \max\Big\{ &\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n-\kappa-\textstyle\sum_{i=1}^n y_i\}|, \\ &\|\max\{0, y_{I_{0+}^*\cup I_{01}^*\cup J} - 1\}\|_\infty,\ \|x_{I_{0+}^*\cup I_{01}^*\cup J}\|_\infty,\ \|\max\{0, -y_{I_{0+}^*\cup I_{01}^*\cup J}\}\|_\infty, \\ &\|y_{I_{\pm 0}^*\cup J^C}\|_\infty \quad \Big\}, \end{aligned}$$

then (1.2) *has a local error bound at* (x^*, y^*) *with*

$$\begin{aligned} r(x,y) := \max\Big\{ &\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n-\kappa-\textstyle\sum_{i=1}^n y_i\}|,\ \|\max\{0, y-1\}\|_\infty, \\ &\|\min\{y, |x|\}\|_\infty \quad \Big\}, \end{aligned} \tag{3.45}$$

where I_{0+}^*, I_{01}^*, $I_{\pm 0}^*$ *are short for* $I_{0+}(x^*, y^*)$, $I_{01}(x^*, y^*)$ *and* $I_{\pm 0}(x^*, y^*)$ *and the operations* max, min *and taking the absolute value on the vectors are understood component-wise.*

Proof. By assumption, for every $J \subseteq I_{00}(x^*, y^*)$, there is a neighbourhood $U((x^*, y^*), J)$ of (x^*, y^*) and a constant $c(J) > 0$ such that

$$\operatorname{dist}_{Z(J)}(x,y) \le c(J) \cdot r_J(x,y) \quad \forall (x,y) \in U((x^*, y^*), J). \tag{3.46}$$

Setting $U((x^*, y^*)) := \bigcap_{J \subseteq I_{00}^*} U((x^*, y^*), J)$ and $C := \max_{J \subseteq I_{00}^*} c(J)$ we obtain

$$\operatorname{dist}_{Z(J)}(x,y) \le C \cdot r_J(x,y) \quad \forall (x,y) \in U((x^*, y^*))\ \forall J \subseteq I_{00}^* \tag{3.47}$$

Since $Z(\hat{J}) \subseteq Z$ for every $\hat{J}$, we have (using (3.47))

$$\operatorname{dist}_Z(x,y) \le \operatorname{dist}_{Z(\hat{J})}(x,y) \le C \cdot r_{\hat{J}}(x,y) \tag{3.48}$$

for each $(x, y) \in U((x^*, y^*))$ and each $\hat{J} \subseteq I_{00}^*$. Furthermore, there is a neighbourhood $\bar{U}(x^*, y^*)$ of (x^*, y^*), such that we have

$$y_i > |x_i| \quad \forall i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*), \tag{3.49}$$

$$y_i < |x_i| \quad \forall i \in I_{\pm 0}(x^*, y^*) \tag{3.50}$$

for all $(x, y) \in \bar{U}(x^*, y^*)$.
Now let $(x, y) \in U((x^*, y^*)) \cap \bar{U}(x^*, y^*)$ be fixed but arbitrary and set

$$\hat{J} := \{i \in I_{00}(x^*, y^*) \,:\, y_i > |x_i|\}. \tag{3.51}$$

Then

$$\hat{J}^C := \{i \in I_{00}(x^*, y^*) \,:\, y_i \le |x_i|\},. \tag{3.52}$$

It follows from (3.49) and (3.50), that

$$|x_i| = \min\{y_i, |x_i|\} \quad \forall i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*) \cup \hat{J}, \tag{3.53}$$

$$y_i = \min\{y_i, |x_i|\} \quad \forall i \in I_{\pm 0}(x^*, y^*) \cup \hat{J}^C. \tag{3.54}$$

Using (3.53) and (3.54) we obtain

$$\begin{aligned}
\max\left\{\|x_{I^*_{0+}\cup I^*_{01}\cup\hat{J}}\|_\infty, \|y_{I^*_{\pm 0}\cup\hat{J}^C}\|_\infty\right\} &= \max\left\{\max_{i\in I^*_{0+}\cup I^*_{01}\cup\hat{J}}\{|x_i|\}, \max_{i\in I^*_{\pm 0}\cup\hat{J}^C}\{|y_i|\}\right\} \\
&= \max\Bigg\{\max_{i\in I^*_{0+}\cup I^*_{01}\cup\hat{J}}\{|\underbrace{\min\{y_i,|x_i|\}}_{\geq 0}|\}, \qquad (3.55) \\
&\qquad \max_{i\in I^*_{\pm 0}\cup\hat{J}^C}\{|\min\{y_i,|x_i|\}|\}\Bigg\} \\
&= \max\left\{\max_{i\in\{1,\dots,n\}}\{|\min\{y_i,|x_i|\}|\}\right\} \\
&= \|\min\{y,|x|\}\|_\infty. \qquad (3.56)
\end{aligned}$$

For $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ and $i \in \{1, \dots, n\}$ distinguish the two following cases:

1. If $y_i < 0$: $|\min\{|x_i|, y_i\}| = |y_i| = |\max\{0, -y_i\}|$.
2. If $y_i \geq 0$: $|\min\{|x_i|, y_i\}| \geq 0 = |\max\{0, -y_i\}|$.

We thus have

$$|\max\{0, -y_i\}| \leq |\min\{|x_i|, y_i\}| \qquad \forall i \in \{1, \dots, n\}. \tag{3.57}$$

Using (3.56) and (3.57), we moreover have

$$\max\left\{\|x_{I^*_{0+}\cup I^*_{01}\cup\hat{J}}\|_\infty,\ \|\max\{0, -y_{I^*_{0+}\cup I^*_{01}\cup\hat{J}}\}\|_\infty,\ \|y_{I^*_{\pm 0}\cup\hat{J}^C}\|_\infty\right\} \leq \|\min\{y,|x|\}\|_\infty. \tag{3.58}$$

Since $I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*) \cup \hat{J} \subseteq \{1, \dots, n\}$, we also have

$$\|\max\{0, y_{I^*_{0+}\cup I^*_{01}\cup\hat{J}} - 1\}\|_\infty \leq \|\max\{0, y - 1\}\|_\infty. \tag{3.59}$$

Using (3.58) and (3.59), it follows from (3.48), with the given index set $\hat{J}$ (see (3.52) and (3.51)), that

$$\begin{aligned}
\operatorname{dist}_Z(x, y) &\leq C \cdot r_{\hat{J}}(x, y) \\
&= C \cdot \max\Big\{\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n - \kappa - \textstyle\sum_{i=1}^n y_i\}|, \\
&\qquad \|\max\{0, y_{I^*_{0+}\cup I^*_{01}\cup\hat{J}} - 1\}\|_\infty,\ \|x_{I^*_{0+}\cup I^*_{01}\cup\hat{J}}\|_\infty, \\
&\qquad \|\max\{0, -y_{I^*_{0+}\cup I^*_{01}\cup\hat{J}}\}\|_\infty,\ \|y_{I^*_{\pm 0}\cup\hat{J}^C}\|_\infty \Big\} \\
&\leq C \cdot \max\Big\{\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n - \kappa - \textstyle\sum_{i=1}^n y_i\}|, \\
&\qquad \|\max\{0, y - 1\}\|_\infty,\ \|\min\{y, |x|\}\|_\infty \Big\} \\
&= C \cdot r(x, y).
\end{aligned}$$

Since $(x, y) \in U((x^*, y^*)) \cap \bar{U}(x^*, y^*)$ was arbitrary, we have

$$\mathrm{dist}_Z(x, y) \leq C \cdot r(x, y)$$

for all $(x, y) \in U((x^*, y^*)) \cap \bar{U}(x^*, y^*)$. It is easy (and brief) to show

$$y_i \geq 0,\ x_i \cdot y_i = 0 \quad \Leftrightarrow \quad \min\{|x_i|, y_i\} = 0$$

for all $i = 1, \dots, n$. Therefore we also have $r(x, y) = 0$ for all $(x, y) \in Z$. □

Lemma 3.67 is essential for the proof of Theorem 3.69, the local error bound for the feasible set of (1.2) under CC-CPLD. For nonlinear programs it is known that CPLD together with local Lipschitz continuity of the gradients of the constraints is sufficient to guarantee the existence of a local error bound. Results such as the following can be found e.g. in [2, Theorem 7] or [64, Theorem 2.1, Corollary 2.1].

Theorem 3.68. *Let $G : \mathbb{R}^N \to \mathbb{R}^M$ and $H : \mathbb{R}^n \to \mathbb{R}^P$ be continuously differentiable,*

$$M := \{z \in \mathbb{R}^n \mid G(z) \leq 0,\ H(z) = 0\}.$$

Let $z^* \in M$ satisfy CPLD and $\nabla G, \nabla H$ be locally Lipschitz continuous around z^*. Then

$$r(z) := \max\{\max\{G_i(z), 0\}\ (i = 1, \dots, M), \quad |H_i(z)|\ (i = 1, \dots, P)\}$$

is a local error bound of M at z^*.

We conclude this section with the following result: The existence of a local error bound for the feasible set of the complementarity formulation. In the proof we use Theorem 3.68 to show that a local error bound holds for the piecewise nonlinear programs.

Theorem 3.69. *Let (x^*, y^*) be feasible for (1.2) and satisfy CC-CPLD. Let ∇g and ∇h be locally Lipschitz continuous around x^*. Then (1.2) has a local error bound at (x^*, y^*) with*

$$r(x, y) = \max\Big\{\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n - \kappa - \sum\nolimits_{i=1}^n y_i\}|,\ \|\max\{0, y - 1\}\|_\infty,\ \|\min\{y, |x|\}\|_\infty \Big\}.$$

In other words: There exists a neighbourhood $N(x^*, y^*)$ of (x^*, y^*) and a constant $c > 0$, such that

$$\mathit{dist}_Z(x, y) \leq c \cdot \max\Big\{\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n - \kappa - \sum\nolimits_{i=1}^n y_i\}|,\ \|\max\{0, y - 1\}\|_\infty,\ \|\min\{y, |x|\}\|_\infty \Big\}$$

for all $(x, y) \in N(x^*, y^*)$.

Proof. Since (x^*, y^*) satisfies CC-CPLD, it follows from Lemma 3.24 that CPLD holds in (x^*, y^*) for NLP(J) for every $J \subseteq I_{00}(x^*, y^*)$. Applying Theorem 3.68 to these programs, it follows that NLP(J) has a local error bound at (x^*, y^*) with

$$r_J(x, y) := \max\Big\{\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n - \kappa - \sum\nolimits_{i=1}^n y_i\}|,\ \|\max\{0, y_{I_{0+}^* \cup I_{01}^* \cup J} - 1\}\|_\infty,\ \|x_{I_{0+}^* \cup I_{01}^* \cup J}\|_\infty,\ \|\max\{0, -y_{I_{0+}^* \cup I_{01}^* \cup J}\}\|_\infty,\ \|y_{I_{\pm 0}^* \cup J}\|_\infty \Big\},$$

for every $J \subseteq I_{00}(x^*, y^*)$. By Lemma 3.67, the program (1.2) consequently has a local error bound at (x^*, y^*) with

$$r(x,y) = \max\Big\{\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n-\kappa-\sum\nolimits_{i=1}^n y_i\}|,\ \|\max\{0, y-1\}\|_\infty,$$
$$\|\min\{y, |x|\}\|_\infty \quad \Big\}.$$

□

The above theorem will help us to prove exactness of a penalty function for (1.2) in Section 4.1.

Concluding Remarks on the Theoretical Results

Due to the structure of the feasible set of the complementarity formulation standard constraint qualifications cannot be expected to hold. We examined this problem in this chapter and reviewed (previously introduced) custom constraint qualifications and first order optimality conditions. Already under the relatively weak CC-GCQ a local minimum is an S-stationary point.

Using the CC-constraint qualifications and ideas from the theory of MPCCs and MPVCs, we were able to derive a number of second order optimality conditions. We derived three main results on second order optimality conditions for the complementarity formulation: Firstly, a second order necessary optimality condition which holds in a local minimum under CC-CRCQ. Secondly, a second order sufficient optimality condition for S-stationary points. Thirdly, for M-stationary points we were able to prove a uniqueness result under a second order condition and CC-CPLD. For the second order sufficient optimality condition and the uniqueness result, in contrast to corresponding results on MPCCs or MPVCs, we consider a smaller subset of the critical cone by excluding all directions (d_x, d_y) with $d_x = 0$. Thereby we take the lack of curvature of the objective function with respect to the auxiliary variable y into account. Additionally, for these three results we provide a corresponding result in terms of the original problem (1.1). With these results we expanded the set of available optimality conditions for the complementarity formulation. The second order conditions can be used to establish convergence results for regularisation methods, which we will consider in the next chapter.

Additionally, in this chapter, we studied optimality conditions for the cardinality constrained problem from [71] and optimality conditions for the case $X = \mathbb{R}^n$ from [6]. Our comparison with the optimality conditions from [6] showed that they are consistent with optimality conditions for the complementarity formulation for this special case. We put the optimality conditions from [71] in relation to the optimality conditions for (1.2) as well. For example, our second order necessary optimality condition for the cardinality constrained problem, Corollary 3.43, captures the second order necessary optimality condition in [71]. The advantage of our optimality conditions for (1.2) is that they can be useful for numerical purposes, as we will see in the following chapter.

Lastly, we used a piecewise decomposition of the complementarity formulation to prove the existence of a local error bound for the complementarity formulation.

In summary, one can say that the transfer of further theoretical results from the theory on standard nonlinear programs to the complementarity formulation under CC-constraint qualifications was successful. To achieve this we applied ideas from the theory on MPCCs and, where required, took the special structure of the complementarity formulation into account. For some results, like the second order optimality conditions, it was also possible to make assertions about the original cardinality constrained problem.

4 Numerical Methods

In this chapter we study numerical methods for the complementarity formulation of cardinality constrained optimization problems. In our setting the complementarity formulation has a smooth objective function and smooth constraint functions. This opens up the possibility to apply a range of methods from nonlinear optimization. Yet, the complementarity formulation poses difficulties from a theoretical point of view. We discussed these difficulties and the means to overcome them in the previous chapter in detail. The numerical methods in this chapter exploit the optimality conditions for the complementarity formulation.
Firstly, we consider two penalisation techniques in Section 4.1: An exact penalty function based on a distance measure for the feasible set of the complementarity formulation, as well as an ℓ^1-norm penalty formulation which is applicable to a special case of (1.2). The results on penalty methods are from the preprint [12].
Secondly, we discuss the application of a sequential quadratic programming (SQP) method to the complementarity formulation in Section 4.2. Motivated by the quadratic subproblems that arise, we consider piecewise nonlinear programs based on the original cardinality constrained problem. We use this decomposition to investigate the behaviour of a SQP method when applied to the complementarity formulation.
Thirdly, we study regularisation methods in Section 4.3. Regularisation methods are a well studied class of methods for mathematical programs with complementarity constraints. Because of the strong link of (1.2) to mathematical programs with complementarity constraints, the adaption of these methods to the complementarity formulation is promising. The results on regularisation methods are from [14, 11, 13].

4.1 Penalisation Techniques

In this section we discuss two penalty approaches for the complementarity formulation. The underlying idea of a penalty method is to replace a constrained optimization problem with an unconstrained optimization problem, with the expectation that this problem will be easier to solve. For a standard nonlinear program of the form (3.1), this would be achieved by considering the problem

$$\min_{x \in \mathbb{R}^n} f(x) + \alpha \cdot r(x) \tag{4.1}$$

for a parameter $\alpha \geq 0$ and a *penalty function* $r : \mathbb{R}^n \to \mathbb{R}$. The *penalty term* $r(x)$ should be zero for feasible points $x \in X$ and positive for infeasible points $x \notin X$. Therefore the objective function of the penalised problem (4.1) accounts for optimality for the original problem, as well as for feasibility. A desirable property, so-called *exactness*, is that a local solution of the original problem is also a local solution of the unconstrained problem.
In Section 4.1.1 we use a distance function to penalise infeasible points. We discuss two variants of this approach: Firstly, we consider an unconstrained penalty problem in the form of (4.1). Secondly, we keep all constraints except for the constraint $x_i \cdot y_i = 0$, which results in a (partially) constrained penalty problem. We add a penalty term to the objective function

which only penalises the complementarity constraint $x_i \cdot y_i = 0$, from which the difficulties regarding (1.2) arise. For these approaches we can show exactness using the local error bound from Chapter 3.
In Section 4.1.2 we consider a special case of (1.2) in which the additional constraint $x \geq 0$ is present. This case includes for instance portfolio selection problems with no short selling. For this case, we can use an ℓ^1-norm penalty term, which is differentiable. Thus we obtain a penalty problem with differentiable objective function and differentiable constraint functions. We discuss the relation between CC-constraint qualifications for the complementarity formulation for the case $x \geq 0$, and (standard) constraint qualifications for the penalised problem. Furthermore, we state relations between M- and S-stationary points and KKT points. Moreover, we show convergence of KKT points of the penalised problem to an M-stationary or S-stationary point respectively.
The results in this section are from [12], which is in preparation.

4.1.1 A Distance-based Penalty Function

In general, a penalty function for (1.2) is of the form

$$(x, y) \mapsto P_\alpha(x, y) = f(x) + \alpha r(x, y)$$

with a penalty parameter $\alpha \geq 0$. As mentioned above, the function $r : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ought to satisfy

$$r(x, y) = 0 \quad \forall (x, y) \in Z \quad \text{and} \quad r(x, y) > 0 \quad \forall (x, y) \notin Z,$$

and is called *penalty term*. Before we begin with the construction of the penalty function, let us state the definition of an exact penalty function.

Definition 4.1. Let (x^*, y^*) be a local minimum of (1.2). A penalty function P_α is called *exact* at (x^*, y^*), if (x^*, y^*) is a local minimum of P_α for all $\alpha \geq 0$ sufficiently large.

Obviously, exactness is a desirable property of a penalty function, because it indeed allows us to replace (1.2) by the unconstrained problem

$$\min_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n} P_\alpha(x, y)$$

with a sufficiently large weight $\alpha \geq 0$. One penalty term, for which exactness in well known, is

$$(x, y) \mapsto \mathrm{dist}_Z(x, y),$$

where the distance is computed with respect to any norm on $\mathbb{R}^n \times \mathbb{R}^n$. The following result can be found e.g. in [19, Proposition 2.4.3].

Theorem 4.2. *Let (x^*, y^*) be feasible for (1.2) and $L \geq 0$ be a local Lipschitz constant of f around x^*. Then*

$$(x, y) \mapsto f(x) + \alpha dist_Z(x, y)$$

is exact for all $\alpha \geq L$.

Note that this result only requires local Lipschitz continuity of the objective f but not necessarily differentiability. Also, the local minimum does not have to be strict. In this thesis, we initially assumed f to be continuously differentiable.

Unfortunately, each evaluation of $\mathrm{dist}_Z(x,y)$ requires the solution of an optimization problem of the type (1.2). Motivated e.g. by [53, 45] we define

$$F(x,y) := \begin{pmatrix} g(x) \\ h(x) \\ e^T y \\ \binom{x_i}{y_i}_{i=1}^n \end{pmatrix} \quad \text{and} \quad \Lambda := (-\infty, 0]^m \times \{0\}^p \times [n-\kappa, \infty) \times C^n,$$

where

$$C := \{(a,b) \in \mathbb{R}^2 \mid b \in [0,1], ab = 0\}.$$

By construction of F and Λ, we have that

$$(x,y) \mapsto r(x,y) := \mathrm{dist}_\Lambda(F(x,y))$$

is a penalty term for (1.2) for any norm on $\mathbb{R}^{m+p+1+2n}$. And contrary to $\mathrm{dist}_Z(x,y)$, we can give a closed form for $\mathrm{dist}_\Lambda(F(x,y))$.

Proposition 4.3. *If $\mathrm{dist}_\Lambda(F(x,y))$ is computed with respect to the ∞-norm, then*

$$\begin{aligned} \mathrm{dist}_\Lambda(F(x,y)) = \max\{ \quad & \max\{g_i(x), 0\} \ (i=1,\dots,m), \quad |h_i(x)| \ (i=1,\dots,p), \\ & \max\{n-\kappa-e^T y, 0\}, \quad |\min\{y_i, \max\{|x_i|, y_i - 1\}\}| \ (i=1,\dots,n)\}. \end{aligned}$$

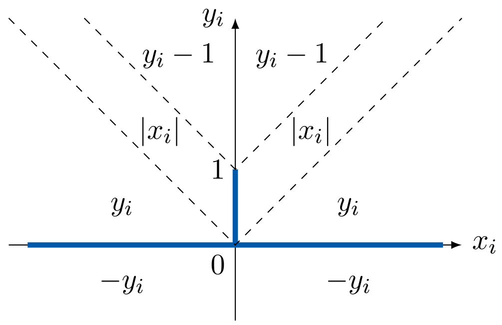


Figure 4.1: Illustration of $\mathrm{dist}_C(x_i, y_i)$

Proof. Due to the Cartesian structure of Λ and the fact that we use the ∞-norm we know

$$\begin{aligned} \mathrm{dist}_\Lambda(F(x,y)) = \max\{ \quad & \mathrm{dist}_{(-\infty,0]}(g_i(x)) \ (i=1,\dots,m), \quad \mathrm{dist}_{\{0\}}(h_i(x)) \ (i=1,\dots,p), \\ & \mathrm{dist}_{[n-\kappa,\infty)}(e^T y), \quad \mathrm{dist}_C(x_i, y_i) \ (i=1,\dots,n) \quad \}. \end{aligned}$$

Here, we immediately obtain

$$\begin{aligned} \mathrm{dist}_{(-\infty,0]}(g_i(x)) &= \max\{g_i(x), 0\}, \\ \mathrm{dist}_{\{0\}}(h_i(x)) &= |h_i(x)|, \\ \mathrm{dist}_{[n-\kappa,\infty)}(e^T y) &= \max\{n-\kappa-e^T y, 0\}. \end{aligned}$$

For the set C, we have

$$
\begin{aligned}
\operatorname{dist}_C(x_i, y_i) &= \begin{cases} |y_i| & \text{if } y_i \le |x_i|, \\ |x_i| & \text{if } y_i - 1 \le |x_i| \le y_i, \\ y_i - 1 & \text{if } y_i - 1 \ge |x_i| \end{cases} \\
&= |\min\{y_i, \max\{|x_i|, y_i - 1\}\}|,
\end{aligned}
$$

see Figure 4.1 for an illustration. $\square$

To show that the penalty function

$$
\begin{aligned}
P_\alpha(x, y) &:= f(x) + \alpha \operatorname{dist}_\Lambda(F(x, y)) \\
&= f(x) + \alpha \max\{\max\{g_i(x), 0\}\ (i = 1, \dots, m), \quad |h_i(x)|\ (i = 1, \dots, p), \\
&\qquad \max\{n - \kappa - e^T y, 0\}, \\
&\qquad |\min\{y_i, \max\{|x_i|, y_i - 1\}\}|\ (i = 1, \dots, n)\}
\end{aligned}
$$

from Proposition 4.3 is exact under suitable assumptions, we use the result on a local error bound for (1.2) from Section 3.5.
To establish the relation between the local error bound from Theorem 3.69 to $\operatorname{dist}_\Lambda(F(x, y))$ we use the following equation. It is straightforward to show that the residual term from Theorem 3.69 is in fact equal to $r(x, y) = \operatorname{dist}_\Lambda(F(x, y))$ as defined in this section. We have

$$
\begin{aligned}
\operatorname{dist}_\Lambda(F(x, y)) &= r(x, y) \\
&= \max\Big\{\|\max\{0, g(x)\}\|_\infty,\ \|h(x)\|_\infty,\ |\max\{0, n - \kappa - \textstyle\sum_{i=1}^n y_i\}|, \\
&\qquad \|\max\{0, y - 1\}\|_\infty,\ \|\min\{y, |x|\}\|_\infty \Big\}.
\end{aligned}
\tag{4.2}
$$

for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$.

Theorem 4.4. *Let (x^*, y^*) be a local minimum of problem (1.2) which satisfies CC-CPLD. Let f, ∇g and ∇h be locally Lipschitz continuous around x^*. Then the penalty function with*

$$
\begin{aligned}
P_\alpha(x, y) &:= f(x) + \alpha\, dist_\Lambda(F(x, y)) \\
&= f(x) + \alpha \max\{\max\{g_i(x), 0\}\ (i = 1, \dots, m), \quad |h_i(x)|\ (i = 1, \dots, p), \\
&\qquad \max\{n - \kappa - e^T y, 0\}, \\
&\qquad |\min\{y_i, \max\{|x_i|, y_i - 1\}\}|\ (i = 1, \dots, n)\}.
\end{aligned}
$$

is exact at (x^*, y^*).

Proof. According to Theorem 3.69, keeping in mind (4.2), there exists a neighbourhood $N(x^*, y^*)$ of (x^*, y^*) such that $\operatorname{dist}_\Lambda(F(x, y))$ is a local error bound of Z with some constant $c > 0$. From Theorem 4.2 we further know that the penalty function $(x, y) \mapsto f(x) + \alpha \operatorname{dist}_Z(x, y)$ is exact at (x^*, y^*) for $\alpha \ge 0$ sufficiently large. This implies for all $(x, y) \in N(x^*, y^*)$

$$
P_\alpha(x, y) = f(x) + \alpha \operatorname{dist}_\Lambda(F(x, y)) \ge f(x) + \frac{\alpha}{c} \operatorname{dist}_Z(x, y) \ge f(x^*)
$$

for $\alpha > 0$ sufficiently large. This shows that (x^*, y^*) is an unconstrained local minimum of P_α and thus P_α is exact at (x^*, y^*). $\square$

Because all norms are equivalent on the finite dimensional vector space $\mathbb{R}^n \times \mathbb{R}^n$, we immediately obtain a class of exact penalty functions for (1.2), analogously to e.g. [36, Satz 5.10].

Corollary 4.5. *Let (x^*, y^*) be a local minimum of (1.2). Let f, ∇g, and ∇h be locally Lipschitz continuous at x^*. If CC-CPLD holds in (x^*, y^*), then all penalty functions of the form*

$$f(x) + \alpha\, dist_\Lambda(F(x, y)),$$

where the distance is computed with respect to an arbitrary norm on $\mathbb{R}^{m+p+1+2n}$ and all penalty functions of the form

$$f(x) + \alpha \| \big(\max\{g(x), 0\},\ |h(x)|,\ \max\{n - \kappa - e^T y, 0\},\ |\min\{y, \max\{|x|, y - e\}\}| \big) \|$$

with an arbitrary norm on $\mathbb{R}^{m+p+1+n}$ are exact at (x^*, y^*).

So far, we have penalised all constraints for the sake of simplicity. For computational purposes it can be advantageous to only penalise the complementarity constraint and keep all other constraints of problem (1.2). Thus consider the partially penalised problem

$$\min_{x,y}\ f(x) + \alpha \| \min\{y, |x|\} \|_\infty \quad \text{s.t.} \quad \begin{aligned} g(x) &\leq 0, & h(x) &= 0, \\ 0 \leq y &\leq e, & e^T y &\geq n - \kappa. \end{aligned} \tag{4.3}$$

Here, we obviously have

$$\min\{y_i, |x_i|\} = 0 \quad \Longleftrightarrow \quad y_i \geq 0,\ x_i y_i = 0.$$

Let $\tilde{Z}$ be the feasible set of (4.3). Then we can show that under the same assumptions as in Theorem 4.4 the penalty function

$$(x, y) \mapsto \tilde{P}_\alpha(x, y) := f(x) + \alpha \| \min\{y, |x|\} \|_\infty$$

is exact at (x^*, y^*). Here, we could again use an arbitrary norm on $\mathbb{R}^n$, but to keep the proof simple, we stick to the ∞-norm.

Corollary 4.6. *Let (x^*, y^*) be a local minimum of problem (1.2) and let CC-CPLD hold in (x^*, y^*). Let $f, \nabla g$, and ∇h be locally Lipschitz-continuous in (x^*, y^*). Then there exists a neighbourhood $N(x^*, y^*)$ of (x^*, y^*) such that (x^*, y^*) is also a local minimum of (4.3) for all $\alpha \geq 0$ sufficiently large, i.e. $\tilde{P}_\alpha$ is exact at (x^*, y^*).*

Proof. Let (x^*, y^*) be a local minimum of problem (1.2) satisfying all assumptions. Then by Theorem 4.4 there exists a neighbourhood $N(x^*, y^*)$ of (x^*, y^*) such that

$$P_\alpha(x, y) = f(x) + \alpha \mathrm{dist}_\Lambda(F(x, y)) \geq P_\alpha(x^*, y^*) \quad \forall (x, y) \in N(x^*, y^*)$$

with $\mathrm{dist}_\Lambda(F(x, y))$ defined in Proposition 4.3.
Now consider an arbitrary point $(x, y) \in \tilde{Z} \cap N(x^*, y^*)$. Then we know

$$g(x) \leq 0, \quad h(x) = 0, \quad e^T y \geq n - \kappa, \quad \text{and} \quad y \leq e$$

and thus obtain

$$\begin{aligned}
\operatorname{dist}_\Lambda(F(x,y)) &= \max\{\max\{g_i(x),0\}\ (i=1,\dots,m), \quad |h_i(x)|\ (i=1,\dots,p), \\
&\qquad \max\{n-\kappa-e^Ty,0\}, \quad |\min\{y_i,\max\{|x_i|,y_i-1\}\}|\ (i=1,\dots,n)\} \\
&= \max\{|\min\{y_i,|x_i|\}|\ (i=1,\dots,n)\} \\
&= \|\min\{y,|x|\}\|_\infty.
\end{aligned}$$

Consequently, we obtain

$$\tilde{P}_\alpha(x,y) = P_\alpha(x,y) \geq P_\alpha(x^*,y^*) = \tilde{P}_\alpha(x^*,y^*) \quad \forall (x,y) \in U \cap \tilde{Z},$$

which shows that $\tilde{P}_\alpha$ is exact at (x^*,y^*). $\square$

Thus, we have provided a reformulation of problem (1.2) both as an unconstrained problem and a standard nonlinear optimization problem, where the used penalty function is exact in both cases.
In Theorem 3.68 and thus in all results based on this theorem, we have to assume that the gradients ∇g and ∇h are locally Lipschitz continuous in order to ensure that CPLD implies the existence of a local error bound. Instead of imposing these additional assumptions on ∇g and ∇h, we can also work with a stronger constraint qualification from the standard theory on nonlinear programs.

Definition 4.7. Let x^* be feasible for (3.1). Then x^* satisfies *pseudonormality* if there exist *no* multipliers $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p$ together with a sequence $(x^k)_k \subseteq \mathbb{R}^n$ converging to x^* such that

- $\sum\limits_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum\limits_{i=1}^{p} \mu_i \nabla h_i(x^*) = 0$,
- $\lambda_i \geq 0$, $\lambda_i g_i(x^*) = 0$ for all $i = 1,\dots,m$,
- $\lambda^T g(x^k) + \mu^T h(x^k) > 0$ for all $k \in \mathbb{N}$.

It directly follows from their definitions that pseudonormality is implied by MFCQ. We will use the following auxiliary result.

Lemma 4.8.
Let (x^*,y^*) be a feasible point of (1.2). If CC-MFCQ holds in (x^*,y^*), then pseudonormality for NLP(J) in (x^*,y^*) holds for all $J \subseteq I_{00}(x^*,y^*)$.

Proof. Using Lemma 3.27, and since MFCQ implies pseudonormality, we only have to consider the remaining case, in which $e^Ty^* = n-\kappa$ and $J \cup I_{0+}(x^*,y^*) = \emptyset$. Since we know that CC-MFCQ holds, we only have to prove that there exist *no* multipliers δ,ν and sequence $(x^k,y^k)_k \to (x^*,y^*)$ such that

- $\delta e = \sum_{i \in I_{01}} \nu_i e_i + \sum_{i \in I_{\pm 0} \cup J^C} \nu_i e_i$,
- $\nu_i \geq 0$ for all $I \in I_{01}(x^*,y^*)$,
- and $\delta(n-\kappa-e^Ty^k) + \sum_{i \in I_{01}} \nu_i(y_i^k - 1) + \sum_{i \in I_{\pm 0} \cup J^C} \nu_i y_i^k > 0$.

However, the first of these conditions implies $\nu_i = \delta$ for all $i \in I_{01}(x^*, y^*) \cup I_{\pm 0}(x^*, y^*) \cup J^C = \{1, \ldots, n\}$ and then due to $|I_{01}(x^*, y^*)| = n - \kappa$ the term in the last condition is zero for any sequence $(y^k)_k$. □

For standard nonlinear programs, it is known that pseudonormality implies the exactness of penalty functions, see e.g. [8, Proposition 4.2] for a proof of the following result under the additional assumption that the local minimum is strict or [53, Theorem 4.5] for an analogous result for mathematical programs with complementarity constraints, that also works for nonstrict local minima.

Theorem 4.9. *Let $G : \mathbb{R}^N \to \mathbb{R}^M$ and $H : \mathbb{R}^n \to \mathbb{R}^P$ be continuously differentiable and*

$$M := \{z \in \mathbb{R}^n \mid G(z) \leq 0,\ H(z) = 0\}.$$

Let $z^* \in M$ satisfy pseudonormality and $F : \mathbb{R}^N \to \mathbb{R}$ be locally Lipschitz continuous around z^*. If z^* is a local minimum of F on M, then

$$F(z) + \alpha \|\big(\max\{G(x), 0\},\ |H(x)|\big)\|_1$$

is an exact penalty function at z^* for all $\alpha > 0$ sufficiently large.

Although the previous result is often formulated with the 1-norm, it also implies exactness of the penalty function with any other norm on $\mathbb{R}^{M+P}$, especially the ∞-norm. We can combine this with the piecewise decomposition to obtain an alternative condition for exactness of the penalty function P_α, which does not require local Lipschitz continuity of ∇g and ∇h.

Theorem 4.10. *Let (x^*, y^*) be a local minimum of problem* (1.2) *and satisfy CC-MFCQ. Let f be locally Lipschitz continuous around x^*. Then the penalty function with*

$$\begin{aligned}
P_\alpha(x, y) \quad &:= \quad f(x) + \alpha dist_\Lambda(F(x, y)) \\
&= \quad f(x) + \alpha \max\{\max\{g_i(x), 0\}\ (i = 1, \ldots, m), \quad |h_i(x)|\ (i = 1, \ldots, p), \\
&\qquad\qquad \max\{n - \kappa - e^T y, 0\}, \\
&\qquad\qquad |\min\{y_i, \max\{|x_i|, y_i - 1\}\}|\ (i = 1, \ldots, n)\}.
\end{aligned}$$

is exact at (x^*, y^*).

Proof. Since (x^*, y^*) is a local minimum of (1.2), it also is a local minimum of NLP(J) for all $J \subseteq I_{00}(x^*, y^*)$. In Lemma 3.27 we have shown that CC-MFCQ at (x^*, y^*) implies pseudonormality for NLP(J) at (x^*, y^*) for all $J \subseteq I_{00}(x^*, y^*)$. Applying Theorem 4.9 with the ∞-norm to all NLP(J), we can find a neighbourhood $N(x^*, y^*)$ of (x^*, y^*) and a constant $\bar{\alpha} \geq 0$ such that for all $J \subseteq I_{00}(x^*, y^*)$ and all $\alpha \geq \bar{\alpha}$ we have

$$f(x) + \alpha r_J(x, y) \geq f(x^*) + \alpha r_J(x^*, y^*) = f(x^*) \quad \forall (x, y) \in N(x^*, y^*).$$

Here $r_J(x, y)$ is defined as in Lemma 3.67. Furthermore, in the proof of Lemma 3.67 we have seen that for every $(x, y) \in U$ we can find a $\hat{J} \subseteq I_{00}(x^*, y^*)$ such that

$$r_{\hat{J}}(x, y) \leq r(x, y) = \mathrm{dist}_\Lambda(F(x, y)),$$

which implies

$$P_\alpha(x, y) = f(x) + \alpha \mathrm{dist}_\Lambda(F(x, y)) \geq f(x) + \alpha r_{\hat{J}}(x, y) \geq f(x^*) = P_\alpha(x^*, y^*).$$

□

Thus, we can prove exactness of the penalty function without additional assumptions on ∇g, ∇h but we have to use a stronger constraint qualification. Using Theorem 4.10 instead of Theorem 4.4 we can also prove Corollaries 4.5 and 4.6 under CC-MFCQ without ∇g and ∇h having to be locally Lipschitz continuous.

4.1.2 Penalty Formulation with ℓ^1-norm Penalty Term

In the previous section, we considered the continuous reformulation (1.2) of the general cardinality constrained optimization problem (1.1). For this problem, we constructed a penalty function, which is exact under suitable assumptions.
In some applications, such as portfolio optimization without short sales, we have the additional constraint $x \geq 0$, and thus consider the reformulated problem

$$\begin{aligned} \min_{x,y} f(x) \quad \text{s.t.} \quad & g(x) \leq 0, && h(x) = 0, \\ & e^T y \geq n - \kappa, && 0 \leq y \leq e, \\ & x \geq 0, \ x \circ y = 0. && \end{aligned} \tag{4.4}$$

Then we end up with a "full" complementarity constraint

$$x_i \geq 0, \quad y_i \in [0,1], \quad x_i \cdot y_i = 0 \quad \forall i = 1, \ldots, n.$$

In this case, we can consider the partially penalised problem

$$\begin{aligned} \min_{x,y} f(x) + \alpha x^T y \quad \text{s.t.} \quad & g(x) \leq 0, && h(x) = 0 \\ & e^T y \geq n - \kappa, && 0 \leq y \leq e, \\ & x \geq 0, && \end{aligned} \tag{4.5}$$

for a parameter $\alpha > 0$. Let Z^p denote the feasible set of the partially penalised problem (4.5). Then we immediately see that $x^T y \geq 0$ for all $(x, y) \in Z^p$ and

$$x^T y = 0 \quad \Longleftrightarrow \quad (x, y) \in Z.$$

Thus, it can make sense to consider $x^T y$ as a penalty term in case $x \geq 0$. In the context of mathematical programs with equilibrium constraints, this was done in [76]. And in contrast to the penalty terms from the previous section, the term $x^T y$ is differentiable. Due to $x \geq 0$, $y \geq 0$ for all $(x, y) \in Z^p$, an alternative interpretation of this penalty term is the observation that

$$\|x \circ y\|_1 = \sum_{i=1}^{n} |x_i y_i| = \sum_{i=1}^{n} x_i y_i = x^T y$$

Thus $x^T y$ can also be seen as the ℓ^1-penalty of the equation $x \circ y = 0$.
However, before we are able to analyse problem (4.5), we have to briefly revisit constraint qualifications and stationarity conditions for problem (4.4). If we apply the constraint qualifications as given in Definition 3.23, the gradients e_i with $i \in I_0(x^*)$ would appear twice, once for the constraint $x \circ y = 0$ and once for the constraint $x \geq 0$. But then of course constraint qualifications such as CC-MFCQ are never satisfied.

For a feasible point (x^*, y^*) of (4.4) we know that $i \notin I_0(x^*)$ implies $x_i^* > 0$. Thus locally we have $x_i > 0$ and consider the tightened program

$$\min_x f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0, \quad x_i = 0 \;\; (i \in I_0(x^*)). \tag{4.6}$$

Although this is exactly the same optimization problem that we used in Section 3.2.1 to define CC-constraint qualifications, it is formally a different tightened program. If we applied the tightened program from Section 3.2.1 to problem (4.4), we would end up with the additional constraint $x \geq 0$.
Based on the tightened problem (4.6), we define constraint qualifications for (4.4) as follows:

Definition 4.11. Let (x^*, y^*) be feasible for (4.4). Then (x^*, y^*) satisfies

(a) *CC-MFCQ (Cardinality Constrained - Mangasarian-Fromovitz Constraint Qualification)* if the gradients

$$\{\nabla g_i(x^*), i \in I_g(x^*)\} \quad \text{and} \quad \{\nabla h_i(x^*), i = 1, \dots, p, \; e_i \; i \in I_0(x^*)\}$$

are positively linearly independent.

(b) *CC-CPLD (Cardinality Constrained - Constant Positive Linear Dependence Constraint Qualification)* if for all subset $I_1 \subseteq I_g(x^*)$, $I_2 \subseteq \{1, \dots, p\}$ and $I_3 \subseteq I_0(x^*)$ such that the gradients

$$\{\nabla g_i(x), i \in I_1\}, \quad \text{and} \quad \{\nabla h_i(x), i \in I_2, \; e_i, i \in I_3\}$$

are positively linearly dependent in $x = x^*$, they remain linearly dependent in a neighbourhood of x^*.

We can also transfer the stationarity concepts by again defining M-stationarity as the KKT conditions of the tightened program (4.6) and S-stationarity as the x-part of the KKT conditions of (4.4). To do so note that for all (x^*, y^*) feasible for problem (4.4) we have

$$I_{\pm 0}(x^*, y^*) = \{i \mid x_i^* \neq 0, y_i^* = 0\} = \{i \mid x_i^* > 0, y_i^* = 0\} =: I_{+0}(x^*, y^*).$$

Then we obtain:

Definition 4.12. A feasible point (x^*, y^*) of (4.4) is called

(i) *M-stationary* (M = Mordukhovich) if there exist multipliers $(\lambda, \mu, \gamma) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$ such that

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \mu_i \nabla h_i(x^*) + \sum_{i=1}^n \gamma_i e_i = 0,& \\ \lambda_i \geq 0, \quad \lambda_i \cdot g_i(x^*) = 0 \quad \forall i = 1, \dots, m,& \\ \gamma_i = 0 \quad \forall i \notin I_0(x^*).& \end{aligned}$$

(ii) *S-stationary* (S = Strong) if it is M-stationary and additionally

$$\gamma_i \leq 0 \quad \forall i \in I_{00}(x^*).$$

Obviously M-stationarity for (4.4) remains unchanged compared to the condition for problem (1.2) because the tightened program is the same. But if we compare the S-stationarity conditions for both problems, we now have the condition $\gamma_i \leq 0$ for all $i \in I_{00}(x^*)$, where before we had the condition $\gamma_i = 0$ for all $i \in I_{00}(x^*)$. In problem (4.4) however, we still have to consider the constraint $x_i \geq 0$ and thus $\gamma_i \leq 0$.
Note that, although we now consider a full complementarity constraint, the constraint qualifications and stationarity conditions still differ from the ones known for mathematical programs with complementarity constraints (MPCC), since they only depend on x. The reason for the different optimality conditions is, as before, that the y-part of the stationarity conditions is trivially satisfied since the objective function does not depend on y and the gradients of all constraints have either an x-component or a y-component but never both. And in points of interest usually $n+1$ constraints on y are active and thus most MPCC constraint qualifications are not satisfied.
Now let us come back to the partially penalised problem (4.5). We have to show that this partially penalised problem is easier to solve than the original problem (4.4) and have to discuss how the solutions or stationary points of both problems are related. We begin by showing that (4.5) inherits MFCQ from the original problem. Using the same arguments, one can also show that it inherits CPLD from the original problem, but since we need MFCQ later, we focus on this constraint qualification only.

Theorem 4.13. *Let (x^*, y^*) be feasible for the original problem (4.4) and satisfy CC-MFCQ.*

(a) Then (x^*, y^*) is feasible for the partially penalised problem (4.5) and MFCQ holds in (x^*, y^*) for (4.5).

(b) There exists a neighbourhood U of (x^*, y^*) such that MFCQ holds at all $(x, y) \in U$ feasible for (4.5).

Proof. We only have to prove part (a). Part (b) then follows as usual from the continuous differentiability of the constraints and the fact that positive linear independence translates to a neighbourhood.
To prove part (a), we have to show that the gradients

$$\begin{aligned}
&\left\{\begin{pmatrix}\nabla g_i(x^*)\\ 0\end{pmatrix}, i \in I_g(x^*), \quad \begin{pmatrix}0\\ -e\end{pmatrix}, \text{if } e^T y^* = n-\kappa, \quad \begin{pmatrix}-e_i\\ 0\end{pmatrix}, i \in I_0(x^*),\right.\\
&\left.\begin{pmatrix}0\\ e_i\end{pmatrix}, i \in I_{01}(x^*, y^*), \quad \begin{pmatrix}0\\ -e_i\end{pmatrix}, i \in I_{00}(x^*, y^*) \cup I_{+0}(x^*, y^*)\right\},\\
&\text{and} \quad \left\{\begin{pmatrix}\nabla h_i(x^*)\\ 0\end{pmatrix}, i = 1, \ldots, p\right\}
\end{aligned}$$

are positively linearly independent. For the gradients with the x-parts, we know this from CC-MFCQ. And the gradients with the y-parts can be positively linearly dependent only if $e^T y^* = n - \kappa$ and $I_{01}(x^*, y^*) = \{1, \ldots, n\}$, which is not possible at the same time. □

This result shows that the partially penalised problem (4.5), contrary to the problem (4.4), does satisfy standard constraint qualifications such as MFCQ. Consequently it is sensible to assume that one can compute KKT points of (4.5). In the subsequent results we therefore want to compare the optimality conditions of both problems.

Theorem 4.14. *Let (x^*, y^*) be feasible point of (4.4). Then we have:*

(a) If (x^*, y^*) is an S-stationary point of (4.4), then there exists a $\hat{\alpha} \geq 0$ such that (x^*, y^*) is a KKT point of (4.5) for all $\alpha \geq \hat{\alpha}$.

(b) If (x^*, y^*) is a KKT point of (4.5), then (x^*, y^*) is an S-stationary point of (4.4).

Proof. (a) Let $(\lambda^*, \mu^*, \gamma^*)$ be multipliers, such that (x^*, y^*) is S-stationary for (1.2).

The KKT conditions for (4.5) in (x^*, y^*), with multipliers $(\lambda, \mu, \xi, \delta, \nu^-, \nu^+) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$, can be expressed as

$$\nabla f(x^*) + \alpha y^* + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) - \sum_{i=1}^{n} \xi_i e_i = 0, \tag{4.7a}$$

$$\sum_{i=1}^{n} \alpha x_i^* e_i + \delta(-e) - \sum_{i=1}^{n} \nu^- e_i + \sum_{i=1}^{n} \nu^+ e_i = 0, \tag{4.7b}$$

$$\lambda_i \geq 0, \quad \lambda_i \cdot g_i(x^*) = 0 \quad \forall i = 1, \ldots, m, \tag{4.7c}$$

$$\xi_i \geq 0, \quad \xi_i \cdot x_i^* = 0 \quad \forall i = 1, \ldots, n, \tag{4.7d}$$

$$\delta \geq 0, \quad \delta \cdot \left(e^T y^* - n + \kappa\right) = 0, \tag{4.7e}$$

$$\nu_i^- \geq 0, \quad \nu_i^- \cdot y_i^* = 0 \quad \forall i = 1, \ldots, n, \tag{4.7f}$$

$$\nu_i^+ \geq 0, \quad \nu_i^+ \cdot (y_i^* - 1) = 0 \quad \forall i = 1, \ldots, n. \tag{4.7g}$$

Here, using the S-stationary multipliers $(\lambda^*, \mu^*, \gamma^*)$ we can choose $(\lambda, \mu, \xi, \delta, \nu^-, \nu^+)$ as follows:

$$\begin{aligned} &\lambda := \lambda^*, \ \mu := \mu^*, \ \delta := 0, \ \nu^+ := 0, \\ &\xi_i := \begin{cases} 0, & \text{if } i \in I_{+0}(x^*, y^*), \\ -\gamma_i^*, & \text{if } i \in I_{00}(x^*, y^*), \\ \alpha y_i^* - \gamma_i^*, & \text{if } i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*), \end{cases} \\ &\nu_i^- := \begin{cases} \alpha x_i^*, & \text{if } i \in I_{+0}(x^*, y^*), \\ 0, & \text{if } i \in \{1, \ldots, n\} \setminus I_{+0}(x^*, y^*). \end{cases} \end{aligned} \tag{4.8}$$

Then we have $\nu_i^- \geq 0$ for all $i = 1, \ldots, n$, since $\alpha \geq 0$ and $x^* \geq 0$. The conditions (4.7a)-(4.7c) and (4.7e)-(4.7g) immediately follow by choice of $(\lambda, \mu, \xi, \delta, \nu^-, \mu^+)$ and because (x^*, y^*) is strongly stationary. Setting

$$\hat{\alpha} := \max \left\{ 0, \max_{i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)} \frac{\gamma_i^*}{y_i^*} \right\},$$

we also have $\xi_i = \alpha y_i^* - \gamma_i^* \geq \hat{\alpha} y_i^* - \gamma_i^* \geq \frac{\gamma_i^*}{y_i^*} \cdot y_i^* - \gamma_i^* = 0$ for all $i \in I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$ and all $\alpha \geq \hat{\alpha}$. For $i \in I_{00}(x^*, y^*)$ we have $\xi_i = -\gamma_i \geq 0$ due to the S-stationarity. Thus (4.7d) also holds and (x^*, y^*) is a KKT point of (4.5) with the chosen multipliers $(\lambda, \mu, \xi, \delta, \nu^-, \nu^+)$.

(b) Recall that we assume (x^*, y^*) to be feasible for (4.4), i.e. (x^*, y^*) is feasible for (4.5) and additionally $(x^*)^T y^* = 0$.

Since (x^*, y^*) is a KKT point of (4.5), there are multipliers $(\lambda, \mu, \xi, \delta, \nu^-, \nu^+)$ such that

$$\nabla f(x^*) + \alpha y^* + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \mu_i \nabla h_i(x^*) - \sum_{i=1}^n \xi_i e_i = 0,$$
$$\lambda_i \geq 0, \quad \lambda_i \cdot g_i(x^*) = 0, \forall i = 1, \dots, m,$$
$$\xi_i \geq 0, \quad \xi_i \cdot x_i^* = 0, \forall i = 1, \dots, n.$$

Setting $\gamma_i := \alpha y_i^* - \xi_i$ for all $i = 1, \dots, n$, we have $\gamma_i = 0$ for all $i \in I_{+0}(x^*, y^*)$ and $\gamma_i = -\xi_i \leq 0$ for all $i \in I_{00}(x^*, y^*)$. Then it follows immediately from the above conditions, that (x^*, y^*) is an S-stationary point of (4.4) with multipliers (λ, μ, γ).

□

In the previous Theorem 4.14(b) we have shown that KKT points of (4.5), which are feasible for (4.4), are S-stationary points of (4.4). The next lemma provides a condition on KKT points of (4.5), which ensures their feasibility for (4.4).

Lemma 4.15. *Let (x^*, y^*) be a KKT point of (4.5), where the multiplier δ corresponding to the constraint $e^T y \geq n - \kappa$ satisfies $\delta = 0$. Then (x^*, y^*) is feasible for (4.4) and S-stationary.*

Proof. We only have to prove that $x_i^* y_i^* = 0$ for all $i = 1, \dots, n$. The S-stationarity then follows from Theorem 4.14(b).
Since (x^*, y^*) is a KKT point of (4.5), there exist multipliers
$(\lambda, \mu, \xi, \delta, \nu^-, \nu^+) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\begin{aligned}
&\alpha x^* + \delta(-e) - \nu^- + \nu^+ = 0, \\
&\delta \geq 0, \quad \delta \cdot \left(e^T y^* - n + \kappa\right) = 0, \\
&\nu_i^- \geq 0, \quad \nu_i^- \cdot y_i^* = 0 \quad \forall i = 1, \dots, n, \\
&\nu_i^+ \geq 0, \quad \nu_i^+ \cdot (y_i^* - 1) = 0 \quad \forall i = 1, \dots, n.
\end{aligned}$$

Here, by assumption $\delta = 0$ and thus for all $i = 1, \dots, n$

$$\alpha x_i^* - \nu_i^- + \nu_i^+ = \delta = 0.$$

Due to the sign constraints on ν_i^+, ν_i^-, this implies that either $x_i^* = 0$ or $x_i^* > 0$ and then $\nu_i^- > 0$, which is only possible if $y_i^* = 0$. Thus, for all $i = 1, \dots, n$ we have $x_i^* y_i^* = 0$. □

Thus, whenever we obtain a KKT point of (4.5) with $\delta = 0$, we know that we have found an S-stationary point of (4.4). The next result considers the case, that we obtain a sequence of – possibly infeasible - KKT points (x^k, y^k) of (4.5) and shows that the limit, if it is feasible, is always M-stationary. In fact, it is even S-stationary if all iterates (x^k, y^k) are infeasible.
The essential observation, that allows to prove this result, goes back to the paper [1] on chance constraints and has been successfully employed in [11] to prove convergence of a Scholtes-type relaxation for cardinality constrained problems.

Theorem 4.16. *Let $\alpha^k \to \infty$ for $k \to \infty$ and $(x^k, y^k)_{k \in \mathbb{N}}$ be a sequence of KKT points of (4.5). Let (x^*, y^*) be a limit point of $(x^k, y^k)_{k \in \mathbb{N}}$, which is feasible for (4.4) and satisfies CC-MFCQ. Then (x^*, y^*) is an M-stationary point of (4.4).*
If the multiplier δ^k corresponding to the constraint $e^T y \geq n - \kappa$ satisfies $\delta^k > 0$ for all $k \in \mathbb{N}$, then (x^*, y^*) is even an S-stationary point of (4.4).

Proof. Since (x^k, y^k) is a KKT point of (4.5), there are multipliers $(\lambda^k, \mu^k, \xi^k, \delta^k, \nu^k) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$ for all $k \in \mathbb{N}$ such that

$$\nabla f(x^k) + \alpha^k y^k + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^n \xi_i^k e_i = 0, \tag{4.9}$$

$$\alpha^k \sum_{i=1}^n x_i^k e_i - \delta^k e + \sum_{i=1}^n \nu_i^k e_i = 0, \tag{4.10}$$

$$\lambda_i^k \geq 0,\ \lambda_i^k \cdot g_i(x^k) = 0, \quad \forall i = 1, \dots, m,$$

$$\xi_i^k \geq 0,\ \xi_i^k \cdot x_i^k = 0, \quad \forall i = 1, \dots, n,$$

$$\delta^k \geq 0,\ \delta^k \cdot (e^T y^k - n + \kappa) = 0,$$

$$\nu_i^k \begin{cases} \leq 0, & \text{if } y_i^k = 0, \\ \geq 0, & \text{if } y_i^k = 1, \\ 0, & \text{otherwise,} \end{cases} \quad \forall i = 1, \dots, n.$$

Define $\gamma_i^k := \alpha^k y_i^k - \xi_i^k$ for all $i = 1, \dots, n$ and all $k \in \mathbb{N}$. From (4.9) we have

$$\nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \nabla \mu_i^k h_i(x^k) + \sum_{i=1}^n \gamma_i^k e_i = 0 \tag{4.11}$$

for all $k \in \mathbb{N}$. We will show by contradiction that the sequence $(\lambda^k, \mu^k, \gamma^k)_{k \in \mathbb{N}}$ is bounded. To this end, assume that $\|(\lambda^k, \mu^k, \gamma^k)\| \to \infty$ as $k \to \infty$. Then we can assume without loss of generality that the sequence

$$\left(\frac{(\lambda^k, \mu^k, \gamma^k)}{\|(\lambda^k, \mu^k, \gamma^k)\|} \right)_{k \in \mathbb{N}}$$

is convergent and

$$\lim_{k \to \infty} \frac{(\lambda^k, \mu^k, \gamma^k)}{\|(\lambda^k, \mu^k, \gamma^k)\|} =: (\bar{\lambda}, \bar{\mu}, \bar{\gamma}) \neq 0.$$

Since $\lambda_i^k \geq 0$ for all $i \in \{1, \dots, m\}$ and $k \in \mathbb{N}$, we have $\bar{\lambda} \geq 0$. For all $i \in \{1, \dots, m\}$ such that $g_i(x^*) < 0$ we have $g_i(x^k) < 0$, hence $\lambda_i^k = 0$ for sufficiently large k, and thus $\bar{\lambda}_i = 0$. We consequently have

$$\operatorname{supp}(\bar{\lambda}) \subseteq I_g(x^*). \tag{4.12}$$

We will show $\operatorname{supp}(\bar{\gamma}) \subseteq I_0(x^*)$ by contradiction. Assume there is an index $j \in \{1, \dots, n\}$ such that $\bar{\gamma}_j \neq 0$ and $x_j^* > 0$. This implies $\gamma_j^k \neq 0$ and $x_j^k > 0$ for sufficiently large $k \in \mathbb{N}$. Using KKT conditions, we thus have $\xi_j^k = 0$ and

$$\alpha^k y_j^k = \gamma_j^k \neq 0 \tag{4.13}$$

for sufficiently large $k \in \mathbb{N}$. Since (x^*, y^*) is feasible for (4.4), we know $y_j^* = 0$. This implies $y_j^k \to 0$ $(k \to \infty)$ and, using (4.13) and the KKT conditions, also $y_j^k > 0$ and $\nu_j^k = 0$ for sufficiently large $k \in \mathbb{N}$. From (4.10) we have $\delta^k = \alpha^k x_j^k > 0$ for sufficiently large $k \in \mathbb{N}$. Because the KKT conditions hold, we have $e^T y^k = n - \kappa$ for sufficiently large $k \in \mathbb{N}$. This means, since $y_j^k \to 0+$ $(k \to \infty)$, we can find an index $l \in \{1, \dots, n\}$ such that y_l^k is monotonically increasing (at least on a subsequence of $(x^k, y^k)_{k \in \mathbb{N}}$) and thus compensating the decrease of y_j^k. We thus have $0 < y_l^k < 1$ and hence $y_l^* > 0$, $x_l^* = 0$ and $\nu_l^k = 0$. Using (4.10) again, we have $\alpha^k x_l^k = \delta^k > 0$, hence $x_l^k > 0$ and $\xi_l^k = 0$, for sufficiently large $k \in \mathbb{N}$. Moreover, we have $\gamma_l^k = \alpha^k y_l^k > 0$ for k sufficiently large. Hence $\bar{\gamma}_l > 0$ and, because $\gamma_j^k \to \bar{\gamma}_j \neq 0$ $(k \to \infty)$ by assumption, we have

$$0 \neq \frac{\bar{\gamma}_j}{\bar{\gamma}_l} = \lim_{k\to\infty} \frac{\gamma_j^k}{\gamma_l^k} = \lim_{k\to\infty} \frac{\alpha^k y_j^k - \xi_j^k}{\alpha^k y_l^k - \xi_l^k} = \lim_{k\to\infty} \frac{y_j^k}{y_l^k} = \frac{y_j^*}{y_l^*} = 0.$$

This is a contradiction. We consequently have shown

$$\operatorname{supp}(\bar{\gamma}) \subseteq I_0(x^*). \tag{4.14}$$

Dividing (4.11) by $\|(\lambda^k, \mu^k, \gamma^k)\|$, taking into account (4.14) and (4.12), and letting $k \to \infty$ we obtain

$$\sum_{i \in I_g(x^*)} \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^{p} \bar{\mu}_i \nabla h_i(x^*) + \sum_{i \in I_0(x^*)} \bar{\gamma}_i e_i = 0.$$

Because $\bar{\lambda} \geq 0$ and $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}) \neq 0$, this is a contradiction to CC-MFCQ in (x^*, y^*). Thus our initial assumption is wrong and the sequence $(\lambda^k, \mu^k, \gamma^k)_{k\in\mathbb{N}}$ is bounded. By choosing a subsequence, if necessary, we can without loss of generality assume that it is convergent. Let $\lim_{k\to\infty}(\lambda^k, \mu^k, \gamma^k) =: (\lambda, \mu, \gamma)$.
From (4.11) we then have

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i=1}^{n} \gamma_i e_i = 0$$

for $k \to \infty$. As above, it can be shown that $\lambda \geq 0$, $\operatorname{supp}(\lambda) \subseteq I_g(x^*)$ and $\operatorname{supp}(\gamma) \subseteq I_0(x^*)$. Consequently (x^*, y^*) is M-stationary for (4.4) with multipliers (λ, μ, γ).
Now let us prove S-stationarity under the additional assumption that $\delta^k > 0$ for all $k \in \mathbb{N}$. Assume that there was an index j with $x_j^* = 0$, $y_j^* = 0$ but

$$\gamma_j = \lim_{k\to\infty} \alpha_k y_j^k - \xi_j^k > 0.$$

This implies $y_j^k > 0$ for all k sufficiently large and since $y_j^k \to 0$ we know $\nu_j^k = 0$ for all k sufficiently large. Using the equation

$$\alpha_k x_j^k + \nu_j^k = \delta^k > 0 \tag{4.15}$$

we thus know $x_j^k > 0$ for k large and therefore $\xi_j^k = 0$.
Since $\delta^k > 0$ implies $e^T y^k = n - \kappa$ for all k but y_j^k is strictly monotonically decreasing (at least on a subsequence). Using a similar argument as before, we know that there exists and index j such that y_l^k is strictly monotonically increasing (at least on a subsequence). This implies $y_l^* > 0$, $\nu_l^k = 0$ and, in view of (4.15), thus $x_j^k > 0$. We hence have $\xi_l^k = 0$ for all k large. Thus, we obtain

$$\frac{\gamma_j}{\gamma_l} = \lim_{k\to\infty} \frac{\alpha^k y_j^k - \xi_j^k}{\alpha^k y_l^k - \xi_l^k} = \lim_{k\to\infty} \frac{y_j^k}{y_j^k} = 0,$$

which is a contradiction to $\gamma_j > 0$. $\square$

Algorithm 1 ℓ^1-Penalty Method

Step 1 Choose $\alpha^0 > 0$, $\sigma > 0$, $(x^0, y^0) \in \mathbb{R}^n \times \mathbb{R}^n$, $k = 1$.

Step 2 Compute a KKT point (x^k, y^k) of (4.5).
If (x^k, y^k) is feasible for (4.4): Stop.

Step 3 Set $\alpha^{k+1} \leftarrow \sigma\alpha^k$, $k \leftarrow k+1$ and go to Step 2.

For a similar result for MPCCs, see [48, Theorem 2.1], one obtains C-stationarity in the limit and has to additionally assume a second order condition to obtain M-stationarity.

Theorem 4.14 and Theorem 4.16 allow us to formulate a penalty scheme for (4.4), given by Algorithm 1. We can proceed as follows: If a computed KKT point of (4.5) is feasible for (4.4), then, by Theorem 4.14, we know that is S-stationary for (4.4) and we stop. By Lemma 4.15, this is the case if $\delta^k = 0$, where δ^k is the multiplier corresponding to the constraint $e^T y \geq n - \kappa$ of (4.5). Otherwise, we have $\delta^k > 0$ for the sequence $(x^k, y^k)_k$. If the sequence converges to some limit (x^*, y^*), and (x^*, y^*) is feasible for (4.4), by Theorem 4.16 we know it is S-stationary, provided it satisfies CC-MFCQ. Therefore, in Step 2 of Algorithm 1 one could check if (x^k, y^k) fulfils the constraints of (4.4) with a tolerance. Otherwise, Algorithm 1 should be stopped after a maximum number of iterations.

4.2 Applying a Sequential Quadratic Programming Method

Since (1.2) is a nonlinear program, one possible approach is to directly apply a method for nonlinear programs, such as the sequential quadratic programming (SQP) method, to it. In Chapter 3 we have discussed that (1.2) cannot be expected to fulfil prerequisites, such as LICQ, for the method's convergence. For mathematical programs with complementarity constraints (MPCCs) a possible approach are smoothing methods, such as studied in [51]. In [30] the direct application of SQP methods to MPCCs was studied with promising results. Motivated by these results, convergence of a piecewise SQP method for MPCCs was studied in [31].

In Chapter 5 we will see that an SQP method applied directly to the complementarity formulation (1.2) can compete with custom methods for some problem instances. Therefore we are going to discuss the direct application of SQP to (1.2) in this section.

For an iterate (x^k, y^k) the quadratic subproblems that occur in the SQP method are given by

$$\begin{aligned}
\min_{d_x, d_y} \quad & \nabla f(x^k)^T d_x + d_x^T H^k d_x + \sum_{i=1}^n \gamma_i (d_x)_i (d_y)_i \\
\text{s.t.} \quad & g_i(x^k) + \nabla g_i(x^k)^T d_x \leq 0, && i = 1, \dots, m, \\
& h_i(x^k) + \nabla h_i(x^k)^T d_x = 0, && i = 1, \dots, p, \\
& e^T y^k + e^T d_y \geq n - \kappa, \\
& 0 \leq y_i^k + e_i^T d_y \leq 1, && i = 1, \dots, n, \\
& x_i^k \cdot y_i^k + x_i^k e_i^T d_y + y^k e_i^T d_x = 0, && i = 1, \dots, n.
\end{aligned} \tag{4.16}$$

Assume that (x^k, y^k) satisfies

$$0 \leq y_i^k \leq 1,\ x_i^k \cdot y_i^k = 0,\ \forall i = 1, \dots, n \quad \text{and} \quad e^T y^k \geq n - \kappa. \tag{4.17}$$

For a start vector for the SQP method, one can choose x^0 such that it has at most κ nonzero components. One could for instance choose y^0 according to $y_i^0 = 1$, for all i with $x_i^0 = 0$, and $y_i^0 = 0$, for all i with $x_i^0 \neq 0$. Then the pair (x^0, y^0) satisfies (4.17). Hence the assumption that (4.17) holds is reasonable. Let us consider the following cases:
Firstly, if $y_i \in (0, 1]$ and $x_i^k = 0$, we have $(d_y)_i \leq 0$, $(d_x)_i = 0$. Hence for the next iterate we have $x_i^{k+1} = x_i^k + (d_x)_i = 0$ and $y_i^{k+1} = y_i^k + (d_y)_i \in [0, 1]$.
Secondly, if $y_i^k = 0$ and $x_i^k \neq 0$, we have $(d_y)_i = 0$. Thus in this case we have $y_i^{k+1} = 0$ and x_i^{k+1} can be arbitrary.
Thirdly, if $y_i^k = x_i^k = 0$, we have $y_i^{k+1} \in [0, 1]$ and x_i^{k+1} can again be arbitrary.
Altogether we have $(d_x)_i = 0$ for all $i \in I_{0+}(x^k, y^k) \cup I_{01}(x^k, y^k)$ and $(d_x)_i$ arbitrary for all $i \in I_{\pm 0}(x^k, y^k) \cup I_{00}(x^k, y^K)$.
Considering the the indexes $i \in I_{00}(x^k, y^k)$, we proceed in a similar way to the derivation of the CC-linearisation cone. Because a feasible point of (1.2) fulfils the complementarity constraint, this should also be the case for the linearisation. Then the last term in the objective function of (4.16) vanishes and, altogether, the quadratic subproblem reduces to

$$\begin{aligned} \min_{d_x} \quad & \nabla f(x^k)^T d_x + d_x^T H^k d_x \\ \text{s.t.} \quad & g_i(x^k) + \nabla g_i(x^k)^T d_x \leq 0, && i = 1, \dots, m, \\ & h_i(x^k) + \nabla h_i(x^k)^T d_x = 0, && i = 1, \dots, p, \\ & (d_x)_i = 0, && i \in I_{0+}(x^k, y^k) \cup I_{01}(x^k, y^k). \end{aligned} \tag{4.18}$$

Like for the definition of the CC-linearisation cone, we set $(d_x)_i \cdot (d_y)_i = 0$ for all $i \in I_{00}(x^k, y^k)$. Then we have $x_i^{k+1} \cdot y_i^{k+1} = 0$ for all $i \in I_{00}(x^k, y^k)$.
In case we have $\|x^k\|_0 = \kappa$, there is a unique y^k such that (4.17) holds, and we have $I_{0+}(x^k, y^k) \cup I_{01}(x^k, y^k) = I_{01}(x^k, y^k) = I_0(x^k)$. In case $\|x^k\|_0 < \kappa$, then the vector y^k is not unique and

$$J^k := I_{0+}(x^k, y^k) \cup I_{01}(x^k, y^k) \subseteq I_0(x^k) \quad \text{with} \quad |J| \geq n - \kappa.$$

Therefore (4.18) is the quadratic subproblem corresponding to a problem of the form

$$\begin{aligned} \min_x f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0, && i = 1, \dots, m, \\ & h_i(x) = 0, && i = 1, \dots, p, \\ & x_i = 0, && i \in J^k, \end{aligned} \tag{4.19}$$

with some $J^k \subseteq I_0(x^k)$ and $|J^k| \geq n - \kappa$. This observation motivates us to consider the following piecewise decomposition of (1.1). For a feasible point x^* of (1.1) let $J \subseteq I_0(x^*)$ such that

$$|J| \geq n - \kappa \tag{4.20}$$

holds. We define the piecewise nonlinear optimization problem

$$
\begin{aligned}
\mathrm{CC}(J): \quad \min_{x\in\mathbb{R}^n} \quad & f(x) && && \\
\text{s.t.} \quad & g_i(x) \le 0, && \forall i = 1,\dots,m, && \text{(4.21a)}\\
& h_i(x) = 0, && \forall i = 1,\dots,p, && \text{(4.21b)}\\
& x_i = 0, && \forall i \in J. && \text{(4.21c)}
\end{aligned}
$$

Condition (4.20) ensures that we have $|I_0(x)| \ge n-\kappa$ for all feasible vectors x of (4.21). Let

$$X(J) := \{x \in X : x_i = 0 \ \forall i \in J\}$$

be the feasible set of CC(J). To study the behaviour of an SQP method applied to (1.2) we will state a piecewise SQP scheme and prove a convergence result using the above piecewise programs. To this end we will apply a Newton iteration to this auxiliary problem for which we expect LICQ to be fulfilled. For mathematical programs with equilibrium constraints similar approaches were used to establish convergence of SQP methods, see [59, 31]. However, we cannot use the piecewise nonlinear programs we used in Chapter 3 for this decomposition. In those piecewise nonlinear programs, with both variables x and y present, LICQ cannot be expected to hold (cf. [15, Section 3]). To overcome this problem we use the above decomposition, which only depends on the variable x. Before we consider the SQP method, we will state some useful relations between CC(J) and (1.1).

Proposition 4.17. *Let x^* be a local solution of* (1.1) *and $J \subseteq I_0(x^*)$, such that* (4.20) *holds. Then x^* is also a local solution for* CC(J).

Proof. Let $J \subseteq I_0(x^*)$, such that (4.20) holds. Since x^* is a local solution of (1.1), we have $x^* \in X$ and $x_i^* = 0$ for all $i \in I_0(x^*) \supseteq J$. Thus x^* is also feasible for (4.21). Furthermore there exists a neighbourhood $N(x^*)$ of x^*, such that $f(x^*) < f(x)$ for all $x \in N(x^*) \cap \{x \in X : \|x\|_0 \le 0\}$. Let $x \in N(x^*) \cap X(J)$. Then $x \in X$ and $x_i = 0$ for all $i \in J$. Because $|J| \ge n-\kappa$, it follows that

$$\|x\|_0 = n - |I_0(x)| \le n - |J| \le n - (n-\kappa) = \kappa.$$

We thus have $x \in N(x^*) \cap \{x \in X : \|x\|_0 \le \kappa\}$ and consequently $f(x^*) < f(x)$. $\square$

The following lemma directly follows from the definition of CC-LICQ.

Lemma 4.18. *Let CC-LICQ hold at x^*. Then LICQ for CC(J) holds at x^* for all $J \subseteq I_0(x^*)$ with $|J| \ge n-\kappa$.*

If x^* is a local minimum of (1.1) and CC-LICQ holds at x^*, we know that (x^*, y) is S-stationary for all y such that $(x^*, y) \in Z$. Moreover, the corresponding multipliers $(\lambda^*, \mu^*, \gamma^*)$ are unique with $\gamma^* = 0$ in case $\|x^*\|_0 < \kappa$, see Proposition 3.34. This motivates the following definition:

Definition 4.19. Let x^* be feasible for (1.1). Then x^* is called *S-stationary*, if there exist multipliers $(\lambda^*, \mu^*, \gamma^*)$ such that

$$
\begin{aligned}
\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla h_i(x^*) + \sum_{i=1}^n \gamma_i^* e_i = 0,&\\
\lambda_i^* \ge 0, \quad \lambda_i^* = 0 \quad \forall i \notin I_g(x^*),&\\
\gamma_i^* = 0 \quad \forall i \notin I_0(x^*).&
\end{aligned}
$$

and either $\|x^*\|_0 = \kappa$ or $\gamma^* = 0$.

For this discussion we will use the term S-stationary for x^* in the above sense. Note that this definition coincides with B-KKT points from [71] (see Definition 3.58). Under CC-LICQ the S-stationary multipliers are unique. Furthermore, if x^* is S-stationary, then (x^*, y) is S-stationary for all y such that $(x^*, y) \in Z$ (in the sense of Definition 3.30). Such S-stationary points are closely related to KKT points of the piecewise problems $\mathrm{CC}(J)$:

Lemma 4.20. *Let x^* be S-stationary and CC-LICQ hold in x^*. Then there exists multipliers $(\lambda^*, \mu^*, \gamma^*) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$ such that $(x^*, \lambda^*, \mu^*, \gamma_J^*)$ is a KKT point of $\mathrm{CC}(J)$ for all $J \subseteq I_0(x^*)$ satisfying (4.20).*
If $\|x^*\|_0 < \kappa$, then $\gamma_J^* = 0$ holds.

Proof. The KKT conditions for $\mathrm{CC}(J)$ in x^* are given by

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i \in J} \gamma_i e_i = 0,$$
$$\lambda_i \geq 0, \;\; \lambda_i \cdot g_i(x^*) = 0 \quad \forall i = 1, \ldots, m.$$

In case $\|x^*\|_0 = \kappa$, then $J = I_0(x^*)$. Hence the above KKT conditions follow from the S-stationary conditions for x^*.
In case $\|x^*\|_0 < \kappa$ it follows from the Definition of an S-stationary point x^* that $\gamma = 0$. Thus the S-stationary conditions are equal to

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) = 0.$$

Hence, in this case, the KKT conditions for $\mathrm{CC}(J)$ in fact hold for any set $J \subseteq \{1, \ldots, n\}$. □

Having the above relations for the piecewise programs $\mathrm{CC}(J)$ at our disposal, we can state the SQP scheme. Algorithm 2 can be seen as a SQP method for each of the piecewise problems $\mathrm{CC}(J)$: In each iteration a subset $J^k \subseteq I_0(x^k)$ such that $|J^k| \geq n - \kappa$ is chosen. Then the quadratic program in Step 4 corresponds to the piecewise problem $\mathrm{CC}(J^k)$. Since $x^k = x^k + d^k$, it follows from constraint (4.24) that the components x_i, $i \in J^k$, are also equal to zero in the next iteration. In case the cardinality constraint is active in x^k, e.g. $|\operatorname{supp}(x^k)| = \kappa$, then we can only choose $J^k = I_0(x^k)$. Otherwise it is possible to choose a subset $J^k \subsetneq I_0(x^*)$ and thus possibly allow a different support in the next iteration.
To prove the local convergence of the piecewise SQP method we will use one of the second order optimality conditions we established in Chapter 3. Therefore, we assume f, g and h to be at least twice continuously differentiable. In Corollary 3.48 we used the critical cone $\mathcal{C}_{\mathcal{X}}(x^*)$ at x^* for (1.1). It is easy to check that

$$\bigcup_{J \subseteq I_0(x^*), |J| \geq n-\kappa} \mathcal{C}_J(x^*) = \mathcal{C}_{\mathcal{X}}(x^*),$$

where $\mathcal{C}_J(x^*)$ is the critical cone for $\mathrm{CC}(J)$ at x^*. Let x^* be S-stationary with multipliers $(\lambda^*, \mu^*, \gamma^*)$ and let CC-LICQ hold at x^*. If additionally

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla^2 h_i(x^*) \right) d_x > 0 \tag{4.25}$$

Algorithm 2 Piecewise SQP Method

Step 1 Choose $x^0 \in \mathbb{R}^n$, $(\lambda^0, \mu^0, \gamma^0) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$ and set $k := 0$.

Step 2 Check criteria to stop algorithm.

Step 3 Choose an index set $J^k \subseteq I_0(x^k)$ with $|J^k| \geq n - \kappa$.

Step 4 Compute a stationary point d^k of $\mathrm{QP}(J^k, x^k, \lambda^k, \mu^k, \gamma^k)$:

$$\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(x^k)^T d + \frac{1}{2} d^T H^k d \\
\text{s.t.} \quad & g_i(x^k) + \nabla g_i(x^k)^T d \leq 0, && \forall i = 1, \ldots, m, && (4.22) \\
& h_i(x^k) + \nabla h_i(x^k)^T d = 0, && \forall i = 1, \ldots, p, && (4.23) \\
& x_i^k + e_i^T d = 0, && \forall i \in J^k. && (4.24)
\end{aligned}$$

with

$$H^k := H(x^k) = \nabla^2 f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla^2 g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla^2 h_i(x^k).$$

Let $(\lambda^{k+1}, \mu^{k+1}, \gamma_J^{k+1})$ be the KKT multipliers of d^k, where the multipliers associated with the constraints (4.22) are λ^{k+1}, the multipliers associated with (4.23) are μ^{k+1} and the multipliers associated with (4.24) are γ_J^{k+1}. Set $\gamma_{(J^k)^C} = 0$.

If the above quadratic program has more than one KKT point, choose the KKT point $(d^k, \lambda^{k+1}, \mu^{k+1}, \gamma_J^{k+1})$, such that the distance

$$\|(x^k + d^k, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1}) - (x^k, \lambda^k, \mu^k, \gamma^k)\|$$

is minimal.

Step 5 Set $x^{k+1} := x^k + d^k$, $k \leftarrow k + 1$ and go to Step 2.

holds for all $d_x \in \bigcup_{J \subseteq I_0(x^*)} \mathcal{C}_J(x^*) \setminus \{0\}$, we have by Corollary 3.48 that x^* is a strict local minimum of (1.1). We will use the following auxiliary result.

Lemma 4.21. *Let x^* be feasible for* (1.1) *such that the following conditions hold.*

(i) The functions f, g and h are twice continuously differentiable,

(ii) CC-LICQ holds in x^*,

(iii) x^* is an S-stationary point with (unique) multipliers $(\lambda^*, \mu^*, \gamma^*)$,

(iv) we have $\lambda_i^* > 0$ for all $i \in I_g(x^*)$ and $\lambda_i^* = 0$ for all $i \in I_g(x^*)^C$ (strict complementarity),

(v) the second order sufficiency condition (4.25) *holds.*

For $J \subseteq I_0(x^*)$ with $|J| \geq n - \kappa$ and $J^C = I_0(x^*) \setminus J$ let further

$$F_J : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n,$$

$$(x, \lambda, \mu, \gamma) \mapsto F_J(x, \lambda, \mu, \gamma) := \begin{pmatrix} \nabla_x L(x, \lambda, \mu, \gamma) \\ \min(-g(x), \lambda) \\ h(x) \\ x_J \\ \gamma_{J^C} \end{pmatrix}.$$

Then there exists an $\varepsilon > 0$ such that F_J is continuously differentiable on $B_\varepsilon(x^*, \lambda^*, \mu^*, \gamma^*)$. Moreover, for any $(x^0, \lambda^0, \mu^0, \gamma^0) \in U_\varepsilon(x^*, \lambda^*, \mu^*, \gamma^*)$ the sequence $(x^k, \lambda^k, \mu^k, \gamma^k)_{k\in\mathbb{N}}$, defined by

$$(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1}) := (x^k, \lambda^k, \mu^k, \gamma^k) + d^k,$$

where d^k solves

$$DF_J(x^k, \lambda^k, \mu^k, \gamma^k)d^k = -F_J(x^k, \lambda^k, \mu^k, \gamma^k), \tag{4.26}$$

i.e. the Newton iterations, converges superlinearly to $(x^*, \lambda^*, \mu^*, \gamma^*)$, and

$$\|(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1}) - (x^*, \lambda^*, \mu^*, \gamma^*)\| \leq \frac{1}{2}\|(x^k, \lambda^k, \mu^k, \gamma^k) - (x^*, \lambda^*, \mu^*, \gamma^*)\|$$

for all $k \in \mathbb{N}$.

Proof. Since x^* is S-stationary and CC-LICQ holds, we also have that x^* is a KKT point of CC(J) and $F_J(x^*, \lambda^*, \mu^*, \gamma^*) = 0$, since either $J^C = \emptyset$ or $\gamma^* = 0$ and therefore $\gamma^*_{J^C} = 0$. Because of the continuity of g_i, $i \in \{1, \ldots, m\}$, and the strict complementarity, there exists an $\tilde{\varepsilon} > 0$, such that for all $(x, \lambda, \mu, \gamma) \in B_{\tilde{\varepsilon}}(x^*, \lambda^*, \mu^*, \gamma^*)$ such that

$$\begin{aligned} -g_i(x) &< \lambda_i \qquad \forall i \in I_g(x^*), \\ -g_i(x) &> \lambda_i \qquad \forall i \in I_g(x^*)^C. \end{aligned}$$

Thus we can write

$$\min(-g(x), \lambda) = (\min(-g_i(x), \lambda_i), \ldots, \min(-g_m(x), \lambda_m)) = \begin{pmatrix} -g(x)_{I_g(x^*)} \\ \lambda_{I_g(x^*)^C} \end{pmatrix}$$

for all $(x, \lambda, \mu, \gamma) \in B_{\tilde{\varepsilon}}(x^*, \lambda^*, \mu^*, \gamma^*)$. We consequently have

$$F_J(x, \lambda, \mu, \gamma) = \begin{pmatrix} \nabla_x L(x, \lambda, \mu, \gamma) \\ -g(x)_{I_g(x^*)} \\ \lambda_{I_g(x^*)^C} \\ h(x) \\ x_J \\ \gamma_{J^C} \end{pmatrix}$$

for all $(x, \lambda, \mu, \gamma) \in B_{\tilde{\varepsilon}}(x^*, \lambda^*, \mu^*, \gamma^*)$ (preserving the order of the above vectors' components). Since the functions f, g and h are twice continuously differentiable the function $\nabla_x L$ is continuously differentiable as well as the functions $(x, \lambda, \mu, \gamma) \mapsto -g(x)_{I_g(x^*)}$, $(x, \lambda, \mu, \gamma) \mapsto$

$\lambda_{I_g(x^*)^C}$, $(x,\lambda,\mu,\gamma)\mapsto x_J$ and $(x,\lambda,\mu,\gamma)\mapsto \gamma_{J^C}$. Consequently F_J is continuously differentiable on $B_{\tilde{\varepsilon}}(x^*,\lambda^*,\mu^*,\gamma^*)$ and

$$DF_J(x,\lambda,\mu,\gamma) = \begin{pmatrix} \nabla^2_{xx}L(x,\lambda,\mu,\gamma) & \nabla g(x) & \nabla h(x) & I_n \\ -\nabla g_{I_g(x^*)}(x)^T & 0 & 0 & 0 \\ 0 & I_{I_g(x^*)^C} & 0 & 0 \\ \nabla h(x)^T & 0 & 0 & 0 \\ I_J & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{J^C} \end{pmatrix}.$$

Now let $DF_J(x^*,\lambda^*,\mu^*,\gamma^*)d = 0$, with $d = (d_x, d_\lambda, d_\mu, d_\gamma) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$, i.e.

$$\nabla^2_{xx}L(x^*,\lambda^*,\mu^*,\gamma^*)d_x + \nabla g(x^*)d_\lambda + \nabla h(x^*)d_\mu + I_n d_\gamma = 0, \tag{4.27a}$$

$$-\nabla g_{I_g}(x^*)^T d_x = 0, \tag{4.27b}$$

$$I_{I_g(x^*)^C} d_\lambda = 0, \tag{4.27c}$$

$$\nabla h(x^*)^T d_x = 0, \tag{4.27d}$$

$$I_J d_x = 0, \tag{4.27e}$$

$$I_{J^C} d_\gamma = 0. \tag{4.27f}$$

Taking the inner product of both sides of (4.27a) with d_x and taking into account (4.27b)-(4.27f) we have

$$d_x^T \nabla^2_{xx}L(x^*,\lambda^*,\mu^*,\gamma^*)d_x = 0 \tag{4.28}$$

From (4.27b), (4.27d) and (4.27e) we have $d_x \in C_J(x^*)$. Since the second order sufficiency condition holds in $(x^*,\lambda^*,\mu^*,\gamma^*)$, we have $d_x = 0$. Thus equation (4.27a) can be simplified to

$$\nabla g(x^*)d_\lambda + \nabla h(x^*)d_\mu + I_n d_\gamma = 0.$$

Using (4.27c) and (4.27f) this can be written as

$$\sum_{i\in I_g(x^*)} (d_\lambda)_i \nabla g_i(x^*) + \sum_{i=1}^{p} (d_\mu)_i \nabla h_i(x^*) + \sum_{i\in J} (d_\gamma)_i e_i = 0$$

Since $J \subseteq I_0(x^*)$ and CC-LICQ holds in x^*, and together with (4.27c) and (4.27f), we have $d_\lambda = 0$, $d_\mu = 0$ and $d_\gamma = 0$. Altogether it follows, that $d = 0$ and hence $DF_J(x^*,\lambda^*,\mu^*,\gamma^*)$ is regular.

From e.g. [36, Satz 5.26] it follows, that there exists an $\varepsilon > 0$, such that the sequence of Newton iterations $(x^k,\lambda^k,\mu^k,\gamma^k)_{k\in\mathbb{N}}$ is well defined and converges superlinearly towards $(x^*,\lambda^*,\mu^*,\gamma^*)$. This result also yields that

$$\|(x^{k+1},\lambda^{k+1},\mu^{k+1},\gamma^{k+1}) - (x^*,\lambda^*,\mu^*,\gamma^*)\| \leq \frac{1}{2}\|(x^k,\lambda^k,\mu^k,\gamma^k) - (x^*,\lambda^*,\mu^*,\gamma^*)\|$$

for all $k \in \mathbb{N}$. $\square$

If $\nabla^2 f$, $\nabla^2 g$ and $\nabla^2 h$ are locally Lipschitz continuous one even obtains quadratic convergence from the above lemma. The following result states that the Newton iteration from the previous Lemma is in fact a KKT point of the quadratic programs in Step 4 of the SQP method.

Lemma 4.22. *Let x^* be feasible for* (1.1) *and let the following conditions hold.*

(i) The functions f, g and h are twice continuously differentiable,

(ii) CC-LICQ holds in x^*,

(iii) x^* is an S-stationary point with (unique) multipliers $(\lambda^*, \mu^*, \gamma^*)$,

(iv) we have $\lambda_i^* > 0$ for all $i \in I_g(x^*)$ and $\lambda_i^* = 0$ for all $i \in I_g(x^*)^C$ (strict complementarity),

(v) the second order condition (4.25) *holds.*

Then there exists a neighbourhood $U(x^*, \lambda^*, \mu^*, \gamma^*)$ of $(x^*, \lambda^*, \mu^*, \gamma^*)$ such that for all vectors $(x^k, \lambda^k, \mu^k, \gamma^k) \in U(x^*, \lambda^*, \mu^*, \gamma^*)$ and all $J^k \subseteq I_0(x^k)$ with $|J^k| \geq n - \kappa$ there exists a unique solution $(d_x^k, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1})$ to the KKT conditions of QP($J^k, x^k, \lambda^k, \mu^k, \gamma^k$) in $U(x^*, \lambda^*, \mu^*, \gamma^*)$.
Furthermore, if we extend this solution by $\gamma_i^{k+1} = 0$ for all $i \notin J^k$, then

$$(d_x^k, d_\lambda^k, d_\mu^k, d_\gamma^k) := (d_x^k, \lambda^{k+1} - \lambda^k, \mu^{k+1} - \mu^k, \gamma^{k+1} - \gamma^k)$$

coincides with the solution of (4.26).

Proof. Let $J^k \subseteq J(x^*)$ and $|J^k| \geq n - \kappa$. For all x^k sufficiently close to x^* we have $I_0(x^*) \supseteq I_0(x^k) \supseteq J^k$. A point $(\tilde{d}, \tilde{\lambda}, \tilde{\mu}, \tilde{\gamma})$ fulfils KKT conditions of QP(J^k) if and only if

$$\begin{aligned}
\nabla f(x^k) + \nabla_{xx}^2 L(x^k, \lambda^k, \mu^k, \gamma^k)\tilde{d} + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(x^k) + \sum_{i=1}^m \tilde{\mu}_i \nabla h_i(x^k) + \sum_{i=1}^n \tilde{\gamma}_i e_i &= 0, \\
\min(-g_i(x^k) - \nabla g_i(x^k)^T \tilde{d}, \tilde{\lambda}_i) &= 0 \ \forall i = 1, \dots, m, \\
h_i(x^k) + \nabla h_i(x^k)^T \tilde{d} &= 0 \ \forall i = 1, \dots, p, \\
x_i^k + e_i^T \tilde{d} &= 0 \ \forall i \in J^k, \\
\tilde{\gamma}_i &= 0 \ \forall i \in (J^k)^C,
\end{aligned}$$

where we added the condition $\gamma_i = 0$ for all $i \notin J^k$ to be able to work with $\gamma \in \mathbb{R}^n$. Setting $\tilde{d} = \tilde{x} - x^k$ these conditions are satisfied, if and only if the point $(\tilde{x}, \tilde{\lambda}, \tilde{\mu}, \tilde{\gamma})$ satisfies

$$\nabla f(x^k) + \nabla_{xx}^2 L(x^k, \lambda^k, \mu^k, \gamma^k)(\tilde{x} - x^k) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(x^k)$$
$$+ \sum_{i=1}^p \tilde{\mu}_i \nabla h_i(x^k) + \sum_{i=1}^n \tilde{\gamma}_i e_i = 0, \tag{4.29a}$$
$$\min(-g_i(x^k) - \nabla g_i(x^k)^T (\tilde{x} - x^k), \tilde{\lambda}_i) = 0 \quad \forall i = 1, \dots, m, \tag{4.29b}$$
$$h_i(x^k) + \nabla h_i(x^k)^T (\tilde{x} - x^k) = 0 \quad \forall i = 1, \dots, p, \tag{4.29c}$$
$$x_i^k + e_i^T (\tilde{x} - x^k) = 0 \quad \forall i \in J^k, \tag{4.29d}$$
$$\tilde{\gamma}_i = 0 \quad \forall i \in (J^k)^C. \tag{4.29e}$$

From the strict complementarity in $(x^*, \lambda^*, \mu^*, \gamma^*)$ and the continuity of the functions g_i and ∇g_i for all $i = 1, \dots, m$, there exists an $\varepsilon_1 > 0$, such that

$$\begin{aligned}
-g_i(x^k) &< \lambda_i^k \quad \forall i \in I(x^*), \\
-g_i(x^k) &> \lambda_i^k \quad \forall i \in I(x^*)^C,
\end{aligned} \tag{4.30}$$

for all $(x^k, \lambda^k, \mu^k, \gamma^k) \in B_{\varepsilon_1}(x^*, \lambda^*, \mu^*, \gamma^*)$ and

$$\begin{aligned} -g_i(x^k) - \nabla g_i(x^k)^T(x - x^k) &< \lambda_i \quad \forall i \in I(x^*), \\ -g_i(x^k) - \nabla g_i(x^k)^T(x - x^k) &> \lambda_i \quad \forall i \in I(x^*)^C, \end{aligned} \tag{4.31}$$

for all $(x^k, \lambda^k, \mu^k, \gamma^k), (x, \lambda, \mu, \gamma) \in B_{3\cdot\varepsilon_1}(x^*, \lambda^*, \mu^*, \gamma^*)$. Choosing $3\varepsilon_1$ for the above neighbourhood helps us with a technical argument at the end of this proof. From Lemma 4.21 it follows that there exists an $\varepsilon_2 > 0$, such that the Newton method for $F_J(x^*, \lambda^*, \mu^*, \gamma^*) = 0$ is well defined for all $(x^0, \lambda^0, \mu^0, \gamma^0) \in U_{\varepsilon_2}(x^*, \lambda^*, \mu^*, \gamma^*)$. Set $\varepsilon := \min\{\varepsilon_2, \varepsilon_1\}$ and let $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1}) \in U_\varepsilon(x^k, \lambda^k, \mu^k, \gamma^k)$ be a an iteration of this Newton-method. Let $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1})$ be the next iteration. From Lemma 4.21 we then also have

$$(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1}) \in U_\varepsilon(x^*, \lambda^*, \mu^*, \gamma^*).$$

Since $\varepsilon \le \varepsilon_1$, it follows from (4.31), that

$$\begin{aligned} -g_i(x^k) - \nabla g_i(x^k)^T(x^{k+1} - x^k) &< \lambda_i^{k+1} \quad \forall i \in I(x^*), \\ -g_i(x^k) - \nabla g_i(x^k)^T(x^{k+1} - x^k) &> \lambda_i^{k+1} \quad \forall i \in I(x^*)^C. \end{aligned} \tag{4.32}$$

The system of linear equations to be solved to compute the Newton iteration, written as

$$DF_J(x^k, \lambda^k, \mu^k, \gamma^k) \begin{pmatrix} x^{k+1} - x^k \\ \lambda^{k+1} - \lambda^k \\ \mu^{k+1} - \mu^k \\ \gamma^{k+1} - \gamma^k \end{pmatrix} = -F_J(x^k, \lambda^k, \mu^k, \gamma^k) \tag{4.33}$$

and, using (4.30), is equivalent to

$$\begin{aligned} \nabla^2_{xx} L(x^k, \lambda^k, \mu^k, \gamma^k)(x^{k+1} - x^k) + Dg(x^k)^T(x^{k+1} - x^k) \\ +Dh(x^k)^T(x^{k+1} - x^k) + I_n(\gamma^{k+1} - \gamma^k) &= -\nabla_x L(x^k, \lambda^k, \mu^k, \gamma^k), \\ (Dg(x^k)^T(x^{k+1} - x^k))_{I_g(x^*)} &= g(x^k)_{I_g(x^*)}, \\ (\lambda^{k+1} - \lambda^k)_{I_g(x^*)^C} &= -\lambda^k_{I_g(x^*)^C}, \\ Dh(x^k)^T(x^{k+1} - x^k) &= -h(x^k), \\ (x^{k+1} - x^k)_J &= -x^k_J, \\ (\gamma^{k+1} - \gamma^k)_{J^C} &= -\gamma^k_{J^C}. \end{aligned}$$

Thus we have

$$\begin{aligned} \nabla^2_{xx} L(x^k, \lambda^k, \mu^k, \gamma^k)(x^{k+1} - x^k) \\ +Dg(x^k)^T(x^{k+1} - x^k) \\ +Dh(x^k)^T(x^{k+1} - x^k) + I_n(\gamma^{k+1} - \gamma^k) &= -\nabla_x L(x^k \lambda^k, \mu^k, \gamma^k), && && \text{(4.34a)} \\ -\nabla g_i(x^k)^T(x^k + 1 - x^k)) &= -g_i(x^k) && \forall i \in I_g(x^*), && \text{(4.34b)} \\ \lambda_i^{k+1} &= 0 && \forall i \in I_g(x^*)^C, && \text{(4.34c)} \\ -\nabla h_i(x^k)^T(x^{k+1} - x^k) &= -h_i(x^k) && \forall i = 1, \dots, p, && \text{(4.34d)} \\ x_i^{k+1} - x_i^k &= -x_i^k && \forall i \in J, && \text{(4.34e)} \\ \gamma_i^{k+1} &= 0 && \forall i \in J^C. && \text{(4.34f)} \end{aligned}$$

From (4.34a) it follows that

$$\nabla^2_{xx} L(x^k, \lambda^k, \mu^k, \gamma^k)(x^{k+1} - x^k) + \sum_{i=1}^{m} \lambda_i^{k+1} \nabla g_i(x^k) + \sum_{i=1}^{p} \mu_i^{k+1} \nabla h_i(x^k) + \sum \gamma_i^{k+1} e_i = -\nabla f(x^k)$$

and thus (4.29a) is satisfied with $(\tilde{x}, \tilde{\lambda}, \tilde{\mu}, \tilde{\gamma}) = (x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1})$. From (4.34d) follows (4.29c), from (4.34e) follows (4.29d) and from (4.34f) follows (4.29e). From (4.32) and (4.34b) we have

$$\min\left(-g_i(x^k) - \nabla g_i(x^k)^T (x^{k+1} - x^k), \lambda_i^{k+1}\right) = -g_i(x^k) - \nabla g_i(x^k)^T (x^{k+1} - x^k) = 0$$

for all $i \in I_g(x^*)$. On the other hand it follows for all $i \in I_g(x^*)^C$ from (4.32) and (4.34c) that

$$\min\left(-g_i(x^k) - \nabla g_i(x^k)^T (x^{k+1} - x^k), \lambda_i^{k+1}\right) = \lambda_i^{k+1} = 0,$$

hence (4.29b) also is fulfilled. Altogether the point $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1})$ satisfies the KKT conditions (4.29a)-(4.29e) of the quadratic program.
It remains to show, that the point $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1})$ is the only KKT point of QP(J^k) in a neighbourhood of $(x^*, \lambda^*, \mu^*, \gamma^*)$ and the closest KKT point to $(x^k, \lambda^k, \mu^k, \gamma^k)$. Because of (4.31), we have

$$\begin{aligned} -g_i(x^k) - \nabla g_i(x^k)^T (\tilde{x}^{k+1} - x^k) &< \tilde{\lambda}_i^{k+1} \quad \forall i \in I(x^*), \\ -g_i(x^k) - \nabla g_i(x^k)^T (\tilde{x}^{k+1} - x^k) &> \tilde{\lambda}_i^{k+1} \quad \forall i \in I(x^*)^C. \end{aligned}$$

From (4.29) and analogously to the previews arguments, it follows that

$$DF_J(x^k, \lambda^k, \mu^k, \gamma^k) \begin{pmatrix} \tilde{x}^{k+1} - x^k \\ \tilde{\lambda}^{k+1} - \lambda^k \\ \tilde{\mu}^{k+1} - \mu^k \\ \tilde{\gamma}^{k+1} - \gamma^k \end{pmatrix} = -F_J(x^k, \lambda^k, \mu^k, \gamma^k).$$

Since $DF_J(x^k, \lambda^k, \mu^k, \gamma^k)$ is regular, we consequently have

$$(\tilde{x}^{k+1}, \tilde{\lambda}^{k+1}, \tilde{\mu}^{k+1}, \tilde{\gamma}^{k+1}) = (x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1}).$$

Thus $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1})$ is the only KKT point of QP(J^k) in $U_{3\cdot\varepsilon_1}(x^*, \lambda^*, \mu^*, \gamma^*)$.
To prove that $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1})$ is indeed the closest KKT point to $(x^k, \lambda^k, \mu^k, \gamma^k)$, let $(\hat{x}^{k+1}, \hat{\lambda}^{k+1}, \hat{\mu}^{k+1}, \hat{\gamma}^{k+1})$ be another KKT point, which is not in $U_{3\cdot\varepsilon_1}(x^*, \lambda^*, \mu^*, \gamma^*)$. Then we have

$$\begin{aligned} &\|(\hat{x}^{k+1}, \hat{\lambda}^{k+1}, \hat{\mu}^{k+1}, \hat{\gamma}^{k+1}) - (x^k, \lambda^k, \mu^k, \gamma^k)\| \\ \geq\ &\underbrace{\|(\hat{x}^{k+1}, \hat{\lambda}^{k+1}, \hat{\mu}^{k+1}, \hat{\gamma}^{k+1}) - (x^*, \lambda^*, \mu^*, \gamma^*)\|}_{\geq 3\varepsilon_1} - \underbrace{\|(x^*, \lambda^*, \mu^*, \gamma^*) - (x^k, \lambda^k, \mu^k, \gamma^k)\|}_{<\varepsilon\leq\varepsilon_1} \\ >\ &2 \cdot \varepsilon_1 \\ \geq\ &2 \cdot \varepsilon \\ >\ &\|(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1}) - (x^*, \lambda^*, \mu^*, \gamma^*)\| + \|(x^*, \lambda^*, \mu^*, \gamma^*) - (x^k, \lambda^k, \mu^k, \gamma^k)\| \\ \geq\ &\|(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \gamma^{k+1}) - (x^k, \lambda^k, \mu^k, \gamma^k)\|. \end{aligned}$$

□

In Step 3 of the piecewise SQP method one can choose the index set J^k and therefore choose from a set of possible quadratic subproblems. The next iterate therefore is computed according to an update rule which is selected from a (finite) set of update rules. If we change the update rules between iterates, the sequence still retains its superlinear convergence.

Lemma 4.23. *Let $x^* \in \mathbb{R}^n$ and for $N \in \mathbb{N}$ let $x \mapsto x^{i,+}$, $i = 1, \ldots, N$, be a (finite) set of update rules satisfying*

$$\forall \varepsilon > 0,\ \forall i = 1, \ldots, N,\ \exists \delta_{\varepsilon,i} > 0 :\ \forall x \in B_{\delta_{\varepsilon,i}}(x^*) :\ \|x^{i,+} - x^*\| \leq \varepsilon \cdot \|x - x^*\|. \tag{4.35}$$

For $x^0 \in \mathbb{R}^n$ consider a sequence $(x^k)_{k\in\mathbb{N}}$ with $x^{k+1} \in \{(x^k)^{i,+} : i = 1, \ldots, N\}$ for $k \in \mathbb{N}$. Then there exists a $\delta > 0$ such that for all $x^0 \in B_\delta(x^*)$ we have $x^k \to x^*$ ($k \to \infty$). Furthermore, for all $\varepsilon > 0$ there exists a $K \in \mathbb{N}$ such that

$$\forall k \geq K : \quad \|x^{k+1} - x^*\| \leq \varepsilon \cdot \|x^k - x^*\|,$$

i.e. (x^k) converges to x^* superlinearly.

Proof. Let $\varepsilon = \frac{1}{2}$ and $\delta = \min_{i=1,\ldots,N} \delta_{\frac{1}{2},i}$. We have $\delta > 0$, since N is a finite number. Let $x \in B_\delta(x^*)$. Then for all $i = 1, \ldots, N$ we have

$$\|x^{i,+} - x^*\| \leq \frac{1}{2} \cdot \|x - x^*\|,$$

and hence $x^{i,+} \in B_\delta(x^*)$. Thus for $x^0 \in B_\delta(x^*)$ the sequence (x^k) also satisfies

$$\|x^{k+1} - x^*\| \leq \frac{1}{2}\|x^k - x^*\|$$

for all $k \in \mathbb{N}$. Thus $x^k \to x^*$.
Now consider an arbitrary $\varepsilon > 0$ and define $\delta_\varepsilon := \min_{i=1,\ldots,N} \delta_{\varepsilon,i}$. Because $x^k \to x^*$ as $k \to \infty$, there exists a $K \in \mathbb{N}$ such that $x^k \in B_{\delta_\varepsilon}(x^*)$ for all $k \geq K$. Since for all $i = 1, \ldots, N$ we have $\delta_\varepsilon \leq \delta_{\varepsilon,i}$, by (4.35) we also have

$$\|x^{k+1} - x^*\| \leq \varepsilon \cdot \|x^k - x^*\| \quad \forall k \geq K.$$

□

Having obtained the previous three lemmas we now are in a position to prove the local convergence of the piecewise SQP method.

Theorem 4.24. *Let CC-LICQ hold in x^* and let x^* be an S-stationary point of* (1.1) *with (unique) multipliers $(\lambda^*, \mu^*, \gamma^*)$. Let f, g and h be twice continuously. Let*

$$\lambda_i^* > 0 \qquad \forall i \in I_g(x^*).$$

and let the second order sufficiency condition (4.25) *hold.*
Then there exists a neighbourhood $U(x^*, \lambda^*, \mu^*, \gamma^*)$ of $(x^*, \lambda^*, \mu^*, \gamma^*)$ such that for all start vectors $(x^0, \lambda^0, \mu^0, \gamma^0) \in U(x^*, \lambda^*, \mu^*, \gamma^*)$ the sequence $(x^k, \lambda^k, \mu^k, \gamma^k)_{k\in\mathbb{N}}$ computed by the piecewise SQP method is well defined and converges superlinearly to $(x^*, \lambda^*, \mu^*, \gamma^*)$.

Proof. Let $J^k \subseteq I_0(x^k)$ be the index set chosen in Step 3 of the piecewise SQP method. For x^k sufficiently close to x^* then also $J^k \subseteq I_0(x^*)$ holds. By Lemma 4.22 there exists a neighbourhood $U(x^*, \lambda^*, \mu^*, \gamma^*)$ of $(x^*, \lambda^*, \mu^*, \gamma^*)$ such that for all $(x^k, \lambda^k, \mu^k, \gamma^k) \in U(x^*, \lambda^*, \mu^*, \gamma^*)$ the solution to the KKT conditions of $\mathrm{QP}(J^k, x^k, \lambda^k, \mu^k, \gamma^k)$ extended by $\gamma_i = 0$, $i \in (J^k)^C$, computed in Step 4 is well defined and corresponds with the Newton iteration for $F_{J^k}(x^*, \lambda^*, \mu^*, \gamma^*) = 0$. By Lemma 4.21 the sequence of Newton iterations converges superlinearly to to $(x^*, \lambda^*, \mu^*, \gamma^*)$. In Step 3 of the piecewise SQP method there is only a finite number of possible index sets J^k to choose from. Choosing $(x^0, \lambda^0, \mu^0, \gamma^0)$ sufficiently close to $(x^*, \lambda^*, \mu^*, \gamma^*)$, it follows from Lemma 4.23 that the sequence $(x^k, \lambda^k, \mu^k, \gamma^k)_{k\in\mathbb{N}}$ converges superlinearly to $(x^*, \lambda^*, \mu^*, \gamma^*)$. □

In Chapter 5 we present numerical results of an SQP solver applied to (1.2). The results indicate that the convergence is fast, however better results can be obtained with regularisation methods. A possible explanation is that, due to (4.24), for the sequence $(x^k)_k$ computed by Algorithm 2 we have $x_i^{k+1} = 0$ for all $i \in J^k$. Hence the sequence can get stuck to a particular support quickly.

4.3 Regularisation Methods

Regularisation methods are a popular class of numerical methods for mathematical programs with complementarity constraints (MPCCs), see [21, 52, 54, 57, 80, 82] for different regularisation methods, [46] for a numerical comparison, and [59, 81] for more information on MPCCs. The similar structure of (1.2) to an MPCC suggest the adaption of these methods. In this section we discuss three regularisation methods for the complementarity-type formulation of cardinality constrained optimization problems: The Kanzow-Schwartz regularisation, a Scholtes-type regularisation, which both have been adapted from the theory of MPCCs, as well as a regularisation which was considered in the context of mathematical programs with chance constraints.

To overcome the difficulties caused by the complementarity constraint the feasible set is relaxed to smooth the nondifferentiable points. To achieve this, new constraints depending on a *regularisation parameter* $t \geq 0$ replace the complementarity constraint. The result is a set $Z(t)$ dependent on the regularisation parameter. Let $(t^k)_{k\in\mathbb{N}}$ be a sequence with $t^k > 0$ for all $k \in \mathbb{N}$ and $t^k \to 0$ as $t \to k$. The new constraints are chosen such that

$$Z \subseteq Z(t) \quad \forall t \geq 0 \quad \text{and} \quad Z(t_2) \subseteq Z(t_1) \quad \forall t_2 \leq t_1.$$

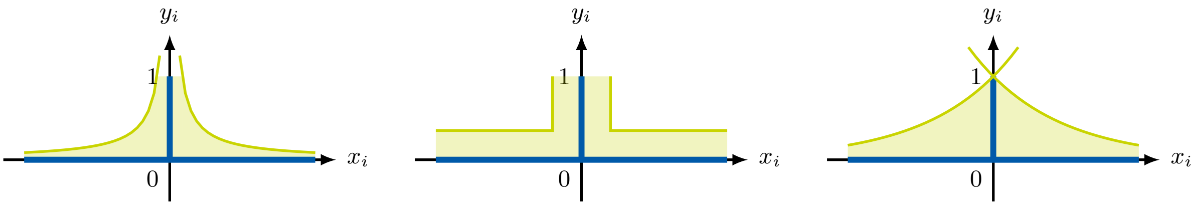
We would like to ensure that for any convergent sequence $(x^k, y^k)_{k\in\mathbb{N}}$, such that $(x^k, y^k) \in Z(t^k)$ for all $k \in \mathbb{N}$, we have $\lim_{k\to\infty}(x^k, y^k) \in Z$. This is fulfilled if

$$Z(t^k) \to Z \quad (k \to \infty)$$

in the sense of Painlevé-Kuratowski (see [77, Chapter 4, Sections A and B]).

Figure 4.2 shows three different ways to relax the complementarity constraint. For the exponential regularisation, for example, the complementarity constraints are replaced by the constraints

$$y_i \leq e^{-t \cdot x_i} \quad \text{and} \quad y_i \leq e^{t \cdot x_i}.$$



(a) Scholtes (b) Kanzow-Schwartz (c) Exponential

Figure 4.2: Pairs (x_i, y_i) that fulfil the complementarity constraint (blue) and the relaxed complementarity constraint (green) for three different regularisations.

for $i = 1, \dots, n$ and $t \geq 0$. We will discuss three possible ways to relax the complementarity constraint, pictured in Figure 4.2, and the regularisation methods arising from them in detail in this section.

Replacing the feasible set Z of (1.2) by the relaxed set $Z(t)$ results in a parametric nonlinear optimization problem. The so-called *regularised problem* NLP(t) is dependent on the regularisation parameter $t \geq 0$.

The underlying idea of the discussed regularisation methods is to compute a sequence of KKT points (x^t, y^t) of NLP(t) for decreasing parameters t. If such a sequence is convergent as $t \to 0$, its limit is feasible for (1.2) by construction of the regularised problems. It is expected that the limit point is an M- or S-stationary point of (1.2). The prerequisites under which this is the case is subject of the convergence analysis of these methods. The functions that describe the constraints which replace the complementarity constraint ought to be differentiable to follow this approach.

The convergence of the sequence (x^t, y^t) is not the only question that arises from following this path. A further question is, whether standard constraint qualifications are fulfilled at feasible points of the regularised problems. In other words, the question is whether NLP(t) has indeed more favourable properties than (1.2). This is of importance if one expects KKT points (x^t, y^t) of NLP(t) to exists. The existence of solutions of NLP(t) in the vicinity of a solution of the complementarity formulation, or in the vicinity of a solution of the cardinality constrained problem, is also desirable. For the Scholtes-type regularisation we address this question using the second order optimality conditions from Section 3.3.

We consider three different regularisation methods for (1.2) in this Section: Firstly, we consider a Scholtes-type regularisation method in Section 4.3.1. It was also originally studied for MPCCs in [80] and adapted to the complementarity formulation in [11]. The Scholtes-type regularisation is illustrated in Figure 4.2a. For this regularisation we discuss the convergence of KKT points of the regularised problems and the validity of standard constraint qualifications for the regularised problem. Moreover, we use the second order optimality conditions from Section 3.3 to extend the convergences theory: We prove the existence of solutions of the regularised problems in the vicinity of a solution to the original cardinality constrained problem (1.1). Given a sequence of KKT points (x^t, y^t) of NLP(t), we further investigate the convergence of x^t to said solution as $t \to 0$. This was originally done in [13].

Secondly, we review the Kanzow-Schwartz regularisation in Section 4.3.2. It was originally introduced for MPCCs [54, 81] and adapted to the complementarity-type formulation of cardinality constrained optimization problems in [14]. This was the first regularisation method for (1.2) for which a convergence result was established. We discuss the convergence of KKT points of the regularised problems as well as the validity of standard constraint qualifications for the regularised problems. Analogously to the Scholtes-type regularisation we extend the convergence theory of this method using second order optimality conditions. Figure 4.2b illustrates the relaxation used by this method.

Lastly, in Section 4.3.3, we consider a regularisation method that uses an exponential function to relax the complementarity constraint. We discuss the convergence of KKT points of the regularised problems as well as the validity of standard constraint qualifications for the regularised problem. This method was studied in the context of mathematical programs with chance constraints in [1]. Figure 4.2c illustrates the relaxation of the complementarity constraint used by this method.

4.3.1 Scholtes-type Regularisation

One of the earliest regularisation methods for MPCCs was the Scholtes regularisation [80]. In the MPCC case, although the theoretical results for the convergence of this method are weaker than those of other regularisation methods, the Scholtes regularisation is numerically very successful [46]. In this section we study a Scholtes-type regularisation for (1.2). The results presented in this section are joint work with Martin Branda, Michal Červinka and Alexandra Schwartz [11, 13].

For the Scholtes-type regularisation, the complementarity constraints are replaced by the constraints

$$-t \le x_i \cdot y_i \le t, \quad i = 1, \dots, n,$$

for a regularisation parameter $t \ge 0$. The relaxation of the complementarity constraints is illustrated in Figure 4.3. The resulting regularised problem is then given by

$$\begin{aligned} \mathrm{NLP}(t): \quad \min_{x,y} \; f(x) \quad \text{s.t.} \quad & g_i(x) \le 0, && \forall i = 1, \dots, m, \\ & h_i(x) = 0, && \forall i = 1, \dots, p, \\ & \textstyle\sum_{i=1}^n y_i \ge n - \kappa, && \\ & -t \le x_i \cdot y_i \le t, && \forall i = 1, \dots, n, \\ & 0 \le y_i \le 1, && \forall i = 1, \dots, n. \end{aligned} \tag{4.36}$$

As before, we denote the feasible set of NLP(t) by $Z(t)$ for $t \ge 0$. We have $Z(t_2) \subseteq Z(t_1)$ for all $0 \le t_2 \le t_1$ as well as $Z \subseteq Z(t)$ for all $t \ge 0$. Like for the other discussed regularisation methods, the approach is to compute a sequence of KKT points of NLP(t) for decreasing parameters $t \to 0$. Any accumulation point of such a sequence is, by construction, feasible for (1.2). Whether it also fulfils a stationarity condition for (1.2) is not obvious. The following theorem gives an answer to this question. We present a slightly modified version compared to [11, Theorem 3.1], which assumes convergence of the x-component of the KKT points only. This modification does not require a change in the proof.

Theorem 4.25. *Let $(t^k)_{k\in\mathbb{N}} \subseteq \mathbb{R}_+$, $t^k \downarrow 0$ ($k \to \infty$) and $(x^k, y^k)_{k\in\mathbb{N}}$ be a sequence of KKT points of* NLP(t^k) *with $x^k \to x^*$ as $k \to \infty$. If CC-MFCQ holds at x^*, then for every accumulation point y^* of the bounded sequence $(y^k)_{k\in\mathbb{N}}$ the pair (x^*, y^*) is an S-stationary point of* (1.2)*.*

Proof. Note that (x^*, y^*) is a feasible point of (1.2). Because (x^k, y^k) is feasible for NLP(t^k) for all $k \in \mathbb{N}$, the sequence $(y^k)_{k\in\mathbb{N}}$ is bounded. Without loss of generality let $y^k \to y^*$

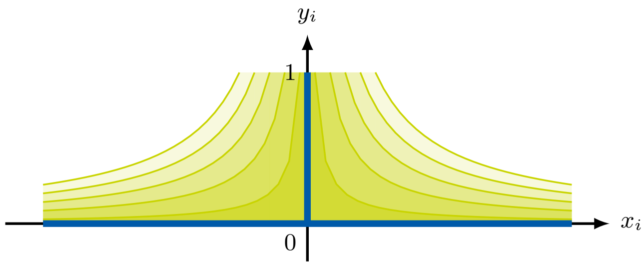


Figure 4.3: Pairs (x_i, y_i) that fulfil the complementarity constraint (blue) and the relaxed complementarity constraint (green) for the Scholtes-type regularisation. The relaxation is illustrated for decreasing regularisation parameters $t_1 > t_2 > \dots$ (darker green for smaller parameters).

as $k \to \infty$. Since $(x^k, y^k)_k$ is a sequence of KKT points of NLP(t^k), there are multipliers $(\lambda^k, \mu^k, \tilde{\gamma}^k, \delta^k, \nu^k)$ for all $k \in \mathbb{N}$ such that

$$\nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) + \sum_{i=1}^n \tilde{\gamma}_i^k y_i^k e_i = 0, \tag{4.37a}$$

$$-\delta^k e + \sum_{i=1}^n \nu_i^k e_i + \sum_{i=1}^n \tilde{\gamma}_i^k x_i^k e_i = 0, \tag{4.37b}$$

$$\lambda_i^k \begin{cases} \geq 0, & \text{if } g_i(x^k) = 0, \\ = 0, & \text{otherwise}, \end{cases} \quad \forall i = 1, \dots, m, \tag{4.37c}$$

$$\delta^k \begin{cases} \geq 0, & \text{if } e^\top y^k = n - \kappa, \\ = 0, & \text{otherwise}, \end{cases} \tag{4.37d}$$

$$\tilde{\gamma}_i^k \begin{cases} \geq 0, & \text{if } x_i^k \cdot y_i^k = t^k, \\ \leq 0, & \text{if } x_i^k \cdot y_i^k = -t^k, \\ = 0, & \text{otherwise}, \end{cases} \quad \forall i = 1, \dots, n, \tag{4.37e}$$

$$\nu_i^k \begin{cases} \leq 0, & \text{if } y_i^k = 0, \\ \geq 0, & \text{if } y_i^k = 1, \\ = 0, & \text{otherwise}, \end{cases} \quad \forall i = 1, \dots, n. \tag{4.37f}$$

Let us first have a closer look at the KKT conditions (4.37). A component-wise inspection of equation (4.37b) yields

$$\delta^k = \nu_i^k + \tilde{\gamma}_i^k x_i^k$$

for all $i = 1, \dots, n$. The sign restrictions on $\tilde{\gamma}^k$ imply $\tilde{\gamma}_i^k \cdot x_i^k \geq 0$. Assuming there is an index $i \in \{1, \dots, n\}$ with $\nu_i^k < 0$, it follows that $y_i^k = 0$ and then, using (4.37e), also $\tilde{\gamma}_i^k = 0$. Thus the above equation yields $0 > \nu_i^k = \delta^k \geq 0$, which is a contradiction. Consequently, we have

$$\nu_i^k \geq 0 \quad \forall i = 1, \dots, n. \tag{4.38}$$

In case $\delta^k > 0$, we have

$$0 < \delta^k = \nu_i^k + \tilde{\gamma}_i^k x_i^k \tag{4.39}$$

for all $i = 1, \dots, n$. Then $\nu_i^k > 0$ or $\tilde{\gamma}_i^k x_i^k > 0$ has to hold for all $i = 1, \dots, n$, which is true only if

$$y_i^k = 1 \quad \text{or} \quad y_i^k = \frac{t^k}{|x_i^k|}. \tag{4.40}$$

For all $k \in \mathbb{N}$ define

$$\gamma_i^k := \tilde{\gamma}_i^k y_i^k \quad \forall i = 1, \dots, n.$$

Boundedness of the multipliers $(\lambda^k, \mu^k, \gamma^k)_k$: We show this by contradiction. Thus, assume that $\lim_{k\to\infty} \|(\lambda^k, \mu^k, \gamma^k)\| = \infty$. Then the sequence

$$\left(\frac{(\lambda^k, \mu^k, \gamma^k)}{\|(\lambda^k, \mu^k, \gamma^k)\|} \right)_{k\in\mathbb{N}}$$

is bounded and without loss of generality let it converge to some limit

$$0 \neq (\bar{\lambda}, \bar{\mu}, \bar{\gamma}) := \lim_{k\to\infty} \frac{(\lambda^k, \mu^k, \gamma^k)}{\|(\lambda^k, \mu^k, \gamma^k)\|}.$$

Clearly, $\bar{\lambda} \geq 0$. Further, for all i such that $g_i(x^*) < 0$ we have $g_i(x^k) < 0$ and thus also $\lambda_i^k = 0$ for all k sufficiently large. That is, we have $\operatorname{supp}(\bar{\lambda}) \subseteq I_g(x^*)$.
Next, to proceed with obtaining a contradiction, let us show that $\operatorname{supp}(\bar{\gamma}) \subseteq I_0(x^*)$. Assume, to the contrary, that there is an index $j \in \{1, \dots, n\}$, such that $x_j^* \neq 0$ and $\bar{\gamma}_j \neq 0$. Then we have $y_j^* = 0$ and consequently

$$x_j^k \neq 0, \quad y_j^k < 1$$

for sufficiently large k. Since $\bar{\gamma}_j^k \neq 0$ we have $\gamma_j^k \neq 0$ and hence $\tilde{\gamma}_j^k \neq 0$ for all k sufficiently large. This implies $\delta^k = \nu_j^k + \tilde{\gamma}_j^k x_j^k > 0$ and thus $\delta^k = \nu_i^k + \tilde{\gamma}_i^k x_i^k > 0$ for all $i = 1, \dots, n$. Due to the KKT conditions, $\delta^k > 0$ is only possible if

$$e^\top y^k = n - \kappa \tag{4.41}$$

for sufficiently large k. Furthermore, for sufficiently large k, $\tilde{\gamma}_j^k \neq 0$ implies

$$0 < y_j^k = \frac{t^k}{|x_j^k|} \quad \text{and} \quad \nu_j^k = 0. \tag{4.42}$$

As $y_j^k \to y_j^* = 0$ and $y_j^k > 0$ hold for k sufficiently large, the sequence $(y_j^k)_k$ is strictly monotonically decreasing (at least on a suitable subsequence). Moreover, since $e^\top y^k = n - \kappa$ for all k sufficiently large (and y^k is a finite-dimensional vector), strict monotone decrease of $(y_j^k)_k$ implies the existence of an index l such that the sequence $(y_l^k)_k$ is strictly monotonically increasing (possibly on a suitable subsequence) and compensates the decrease of $(y_j^k)_k$ in such a way that $e^\top y^k = n - \kappa$ is preserved. Thus, we have

$$y_l^* > 0,\ x_l^* = 0 \quad \text{and} \quad y_l^k < 1,\ \nu_l^k = 0 \text{ for sufficiently large } k.$$

Invoking (4.40) and $y_l^k < 1$ for sufficiently large k, we have

$$y_l^k = \frac{t^k}{|x_l^k|}. \tag{4.43}$$

Since $\nu_j^k = \nu_l^k = 0$, (4.39), (4.42) and (4.43) implies

$$\frac{|\gamma_j^k|}{|\gamma_l^k|} = \frac{|\tilde{\gamma}_j^k \cdot y_j^k|}{|\tilde{\gamma}_l^k \cdot y_l^k|} = \frac{\left|\frac{\delta^k}{x_j^k} \cdot \frac{t^k}{|x_j^k|}\right|}{\left|\frac{\delta^k}{x_l^k} \cdot \frac{t^k}{|x_l^k|}\right|} = \left(\frac{x_l^k}{x_j^k}\right)^2 \xrightarrow[k\to\infty]{} \left(\frac{x_l^*}{x_j^*}\right)^2 = 0.$$

This leads to the contradiction

$$0 \neq |\bar{\gamma}_j| = \lim_{k\to\infty} \frac{|\gamma_j^k|}{\|(\lambda^k, \mu^k, \gamma^k)\|} \leq \lim_{k\to\infty} \frac{|\gamma_j^k|}{|\gamma_l^k|} = 0,$$

which concludes the proof of $\mathrm{supp}(\bar{\gamma}) \subseteq I_0(x^*)$.
Now, dividing (4.37a) by $\|(\lambda^k, \mu^k, \gamma^k)\|$ and taking the limit $k \to \infty$ yields

$$\sum_{i\in I_g(x^*)} \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^{p} \bar{\mu}_i \nabla h_i(x^*) + \sum_{i\in I_0(x^*)} \bar{\gamma}_i e_i = 0.$$

However, this, together with $\bar{\lambda} \geq 0$, $\mathrm{supp}(\bar{\lambda}) \subseteq I_g(x^*)$, $\mathrm{supp}(\bar{\gamma}) \subseteq I_0(x^*)$, and $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}) \neq 0$, is in contradiction with the assumption of CC-MFCQ at (x^*, y^*). Thus, the sequence of multipliers $(\lambda^k, \mu^k, \gamma^k)_k$ is bounded and without loss of generality we can assume that the whole sequence $(\lambda^k, \mu^k, \gamma^k)_k$ converges to some limit

$$(\lambda^*, \mu^*, \gamma^*) := \lim_{k\to\infty} (\lambda^k, \mu^k, \gamma^k).$$

Taking the limit in (4.37a) as $k \to \infty$, we obtain

$$\nabla f(x^*) + \sum_{i\in I_g(x^*)} \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla h_i(x^*) + \sum_{i=1}^{n} \gamma_i^* e_i = 0.$$

S-Stationarity of x^* together with the multipliers $(\lambda^*, \mu^*, \gamma^*)$: Using analogous arguments as previously, we have $\lambda^* \geq 0$ and $\mathrm{supp}(\lambda^*) \subseteq I_g(x^*)$.
To complete the proof, it remains to show that $y_i^* = 0$ implies $\gamma_i^* = 0$. Again, to the contrary, assume that there exists an index j such that $y_j^* = 0$ and $\gamma_j^* \neq 0$. This implies $\gamma_j^k = \tilde{\gamma}_j^k y_j^k \neq 0$ for all k sufficiently large and therefore,

$$0 < y_j^k = \frac{t^k}{|x_j^k|}.$$

Hence the sequence $(y_j^k)_k$ is strictly monotonically decreasing to zero (at least on a suitable subsequence). Thus, we have $x_j^k \neq 0$ and $\nu_j^k = 0$ for all k sufficiently large which together with $\gamma_j^k \neq 0$ implies $\delta^k = \tilde{\gamma}_j^k x_j^k > 0$ and $e^\top y^k = n - \kappa$ for all k sufficiently large. Analogously to the previous part of the proof, there has to exist an index l such that $(y_l^k)_k$ is strictly increasing and

$$0 < y_l^k = \frac{t^k}{|x_l^k|}.$$

This implies $y_l^k \to y_l^* > 0$ and $\nu_l^k = 0$ and thus $\delta^k = \gamma_l^k x_l^k$ and $x_l^k \neq 0$ for all k sufficiently large. Finally, this together with $\gamma_i^k = \tilde{\gamma}_i^k y_i^k$ for all i yields

$$\frac{|\gamma_j^*|}{|\gamma_l^*|} = \lim_{k\to\infty} \frac{|\gamma_j^k|}{|\gamma_l^k|} = \lim_{k\to\infty} \frac{|\tilde{\gamma}_j^k y_j^k|}{|\tilde{\gamma}_l^k y_l^k|} = \lim_{k\to\infty} \frac{\left|\frac{\delta^k}{|x_j^k|} t^k y_j^k\right|}{\left|\frac{\delta^k}{|x_l^k|} t^k y_l^k\right|} = \lim_{k\to\infty} \left(\frac{y_j^k}{y_l^k}\right)^2 = \left(\frac{y_j^*}{y_l^*}\right)^2 = 0,$$

which is a contradiction to $\gamma_j \neq 0$. Consequently (x^*, y^*) is an S-stationary point with multipliers (λ, μ, γ). This completes the proof. $\square$

The above result is unexpected: In the MPCC case the limit of a sequence of KKT points of the regularised problem, under an MFCQ-type constraint qualification, is only C-stationary [46]. However C-stationary points for MPCCs correspond to M-stationary points for (1.2), see Section 3.2.2. Yet, Theorem 4.25 yields S-stationarity of the limit – a stronger result compared to the result for MPCCs.

This is even more surprising, because the adaption of the Kanzow-Schwartz regularisation to (1.2) retains its convergence properties [14]. We will review those results in the following Section 4.3.2. However, the convergence result for the Kanzow-Schwartz regularisation requires the weaker CC-CPLD constraint qualification to hold.

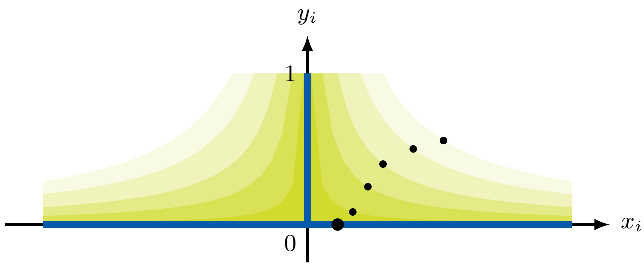


Figure 4.4: KKT points of NLP(t) converge to an S-stationary point of (1.2) as $t \to 0$, illustrated with the complementarity constraints.

In the proof of Theorem 4.25 we exploited the fact that $e^T y^k = n - \kappa$. This idea was used before in [1], where a similar structure is used to reformulate chance constrained optimization problems. We used a slightly different technique than typically used in the MPCC case. Because the required CC-MFCQ (like the other CC-constraint qualifications) depends only on the gradients of the constraints with respect to the variable x, we only normalised the multipliers $(\lambda^k, \mu^k, \gamma^k)$ corresponding to constraints on x. Therefore the verification of $\operatorname{supp}(\bar{\gamma}) \subseteq I_0(x^*)$ in the above proof is more involved.

Instead of normalising only the multipliers $(\lambda^k, \mu^k, \gamma^k)$ one can also normalise all multipliers $(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k)$. This simplifies the verification of the correct support of γ^k but makes it harder to obtain a contradiction to CC-MFCQ. We follow this route in the proof of the following result on the validity of MFCQ for the regularised problems NLP(t).

Theorem 4.26. *Let (x^*, y^*) be feasible for (1.2) and CC-MFCQ hold there. Then there exists a radius $r > 0$ and a constant $T > 0$ such that for all $t \in (0, T]$ standard MFCQ for NLP(t) holds at every $(x, y) \in Z(t)$ with $x \in B_r(x^*)$.*

Proof. Let us assume that the claim is false. Then there exist sequences $(x^k, y^k)_{k\in\mathbb{N}}$ and $(t^k)_k > 0$ such that (x^k, y^k) is feasible for NLP(t^k), $x^k \to x^*$, but MFCQ is violated. Because

(x^k, y^k) is feasible for NLP(t^k) for all $k \in \mathbb{N}$, the sequence $(y^k)_{k\in\mathbb{N}}$ is bounded. Without loss of generality let $y^k \to y^*$ as $k \to \infty$. Because at every (x^k, y^k) MFCQ is violated, we can find multipliers $(\lambda^k, \mu^k, \tilde{\gamma}^k, \delta^k, \nu^k)_k$ such that for all $k \in \mathbb{N}$

$$(\lambda^k, \mu^k, \tilde{\gamma}^k, \delta^k, \nu^k) \neq 0 \quad \forall k \in \mathbb{N}, \tag{4.44}$$

along with condition

$$\sum_{i=1}^{m} \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla h_i(x^k) + \sum_{i=1}^{n} \tilde{\gamma}_i^k y_i^k e_i = 0, \tag{4.45}$$

and (4.37b)-(4.37f) are satisfied. This means the relevant gradients are positively linearly dependent at (x^k, y^k), hence MFCQ is violated. Analogously to the proof of Theorem 4.25, we can show that $\nu_i^k < 0$ cannot occur, thus we have only to consider $\nu_i^k \geq 0$.
Now, let us define $\gamma_i^k := \tilde{\gamma}_i^k y_i^k$ for all $i = 1, \ldots, n$ and all $k \in \mathbb{N}$. Because for all $i = 1, \ldots, n$ and all $k \in \mathbb{N}$ we have $\tilde{\gamma}_i^k \neq 0$ only if $y_i^k \neq 0$, we immediately obtain

$$\operatorname{supp}(\gamma^k) = \operatorname{supp}(\tilde{\gamma}^k) \quad \forall k \in \mathbb{N}.$$

Using (4.37e) we have

$$\tilde{\gamma}_i^k x_i^k = \begin{cases} \gamma_i^k \cdot \frac{x_i^k}{y_i^k}, & \text{if } |x_i^k \cdot y_i^k| = t_k, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, we can write equations (4.45) and (4.37b) for all $k \in \mathbb{N}$ as

$$\sum_{i=1}^{m} \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla h_i(x^k) + \sum_{i=1}^{n} \gamma_i^k e_i = 0, \tag{4.46}$$

$$\delta^k = \begin{cases} \nu_i^k + \gamma_i^k \frac{x_i^k}{y_i^k} & \text{if } |x_i^k \cdot y_i^k| = t_k, \\ \nu_i^k & \text{otherwise.} \end{cases} \tag{4.47}$$

Due to the sign restrictions in (4.37e), we know $\tilde{\gamma}_i^k x_i^k \geq 0$ for all $i = 1, \ldots, n$ and all $k \in \mathbb{N}$. Hence we also have

$$\gamma_i^k x_i^k = \tilde{\gamma}_i^k x_i^k y_i^k \geq 0 \quad \forall i = 1, \ldots, n, \ \forall k \in \mathbb{N}. \tag{4.48}$$

We proceed to deduce a contradiction with CC-MFCQ at (x^*, y^*). Since by assumption $(\lambda^k, \mu^k, \tilde{\gamma}^k, \delta^k, \nu^k) \neq 0$ for all $k \in \mathbb{N}$, we can choose the multipliers without loss of generality such that $\|(\lambda^k, \mu^k, \gamma^k, \delta^k, \nu^k)\| = 1$ for all $k \in \mathbb{N}$ and that the whole sequence converges to some limit

$$0 \neq (\lambda, \mu, \gamma, \delta, \nu) := \lim_{k\to\infty} (\lambda^k, \mu^k, \gamma^k, \delta^k, \nu^k).$$

We have $\lambda \geq 0$. Since for all i such that $g_i(x^*) < 0$ we know $g_i(x^k) < 0$ and thus $\lambda_i^k = 0$ for all k sufficiently large, we have

$$\operatorname{supp}(\lambda) \subseteq I_g(x^*). \tag{4.49}$$

We will prove $\operatorname{supp}(\gamma) \subseteq I_0(x^*)$ by contradiction. To this end we assume that there is an index $j \in \{1, \ldots, n\}$ such that $\gamma_j \neq 0$ and $x_j^* \neq 0$. This implies $|x_i^k \cdot y_i^k| = t_k$ for sufficiently large k and $y_j^* = 0$. Therefore we know $x_j^k \neq 0$ and $y_j^k > 0$ for k sufficiently large and $y_j^k \to 0$ for $k \to \infty$. Thus we have $y_j^k < 1$ and hence $\nu_j^k = 0$ for sufficiently large k. Keeping in mind (4.48) it follows from the j-th row of (4.47) that

$$\delta^k = \nu_j^k + \gamma_j^k \frac{x_j^k}{y_j^k} = \gamma_j^k \frac{x_j^k}{y_j^k} \longrightarrow \infty \quad (k \to \infty).$$

Because $(\lambda^k, \mu^k, \gamma^k, \delta^k, \nu^k)_k$ is convergent, this is a contradiction. Consequently we have

$$\mathrm{supp}(\gamma) \subseteq I_0(x^*). \tag{4.50}$$

It remains to show that $(\lambda, \mu, \gamma) \neq 0$. We will show this also by contradiction. Assume that $(\lambda, \mu, \gamma) = 0$. Since $(\lambda, \mu, \gamma, \delta, \nu) \neq 0$ this implies $(\delta, \nu) \neq 0$. Due to $\nu^k \geq 0$ and (4.48) combined with (4.47) we know $\delta^k \geq \max_{i=1,\dots,n} \nu_i^k$ and thus $(\delta, \nu) \neq 0$ implies $\delta > 0$ and $\delta^k > 0$ for all k sufficiently large. This is only possible, if $e^T y^k = n - \kappa$ for all k large. For all i with $y_i^* > 0$ we know $x_i^* = 0$ and thus $\gamma = 0$ implies

$$0 < \delta^* = \lim_{k \to \infty} \nu_i^k + \gamma_i^k \frac{x_i^k}{y_i^k} = \nu_i.$$

Hence, for all i such that $y_i^* > 0$ we know $y_i^k = 1$ for all k sufficiently large and therefore $y_i^* = 1$. Due to $e^T y^k = n - \kappa < n$ for all k large there exists at least one index j such that $y_j^k = 0$ for all k large and consequently $\nu_j^k = 0$ and $|x_j^k \cdot y_j^k| \neq t_k$. This, however, implies $\delta^k = 0$, a contradiction. Thus the assumption $(\lambda, \mu, \gamma) = 0$ is false and we have

$$(\lambda, \mu, \gamma) \neq 0.$$

Using (4.49) and (4.50), it follows from (4.46) for $k \to \infty$ that

$$\sum_{i \in I_g(x^*)} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i \in I_0(x^*)} \gamma_i e_i = 0. \tag{4.51}$$

Since $(\lambda, \mu, \gamma) \neq 0$ and $\lambda \geq 0$, this is a contradiction to CC-MFCQ. $\square$

We presented a slightly different form of [11, Theorem 3.2]. Similar to Theorem 4.25, it suffices that the x-component of a point $(x, y) \in Z(t)$ is in a neighbourhood of x^*. Theorem 4.26 is similar to the result for the Scholtes regularisation for MPCCs. In that case, one can as well prove validity of MFCQ for the regularised problems locally, provided a MFCQ-type constraint qualification holds.

The above result shows that the regularised problems NLP(t) have indeed better properties than the complementarity formulation. Our next step is to investigate under which conditions the regularised programs NLP(t) posses a local solution in the vicinity of a local solution x^* of (1.1). This in case the regularised problems have local solutions close to a point $(x^*, y^*) \in Z$, these local solutions are KKT points if CC-MFCQ holds at (x^*, y^*). Therefore the next result will be useful to study conditions under which the regularisation method is locally well defined.

Theorem 4.27. *(a) Let x^* be a strict local minimum of* (1.1)*. Then there exist $r > 0$ and $T > 0$ such that for all $t \in (0, T]$ the regularised program NLP(t) has a local minimum (x^t, y^t) with $x^t \in B_r(x^*)$.*

(b) Let (x^*, y^*) be a strict local minimum of (1.2) *with respect to x and $\|x^*\|_0 = \kappa$. Then there exist $r > 0$ and $T > 0$ such that for all $t \in (0, T]$ the regularised program NLP(t) has a local minimum (x^t, y^t) with $x^t \in B_r(x^*)$.*

Proof. (a) By assumption there exists a radius $r > 0$ such that for all $x \in \overline{B_r(x^*)} \setminus \{x^*\}$ feasible for (1.1) we have $f(x) > f(x^*)$.
Now assume that there is no $T > 0$ such that NLP(t) has a local minimum in $Z(t) \cap (B_r(x^*) \times \mathbb{R}^n)$ for all $t \in (0, T]$. Then we can find a sequence $t_k \downarrow 0$ such that NLP(t_k) has no local minimum on $Z(t_k) \cap (B_r(x^*) \times \mathbb{R}^n)$. Since the set $Z(t_k) \cap (\overline{B_r(x^*)} \times \mathbb{R}^n)$ is nonempty and compact (recall that the y-variables are always bounded), f attains a global minimum (x^k, y^k) there. Consequently $x^k \in \partial B_r(x^*)$ and $f(x^k) < f(x^*)$. If we had $f(x^k) \geq f(x^*)$, then the point (x^*, y^*), where $y_i^* = 0$, for all $i \in \operatorname{supp}(x^*)$ and $y_i^* = 1$ for all $i \in I_0(x^*)$, would be a local minimum of f on $Z(t_k) \cap (B_r(x^*) \times \mathbb{R}^n)$.
Since $\partial B_r(x^*)$ is compact, we may assume without loss of generality that $(x^k)_k$ converges to some limit $\bar{x} \in \partial B_r(x^*)$, which implies $\bar{x} \neq x^*$. And since $(y^k)_k$ is bounded, it is also convergent (at least on a subsequence) to some limit $\bar{y}$. Letting $t_k \downarrow 0$, we obtain $(\bar{x}, \bar{y}) \in Z$. Hence $\bar{x}$ is feasible for (1.1). Due to $\bar{x} \neq x^*$ and the choice of r, this yields the contradiction

$$f(x^*) \geq \lim_{k \to \infty} f(x^k) = f(\bar{x}) > f(x^*).$$

(b) We only have to show that the assumptions imply that x^* is a strict local minimum of (1.1). To this end consider an arbitrary sequence $x^k \to x^*$ feasible for (1.1) with $x^k \neq x^*$. Because x^k is feasible for (1.1), the active cardinality constraint $\|x^*\|_0 = \kappa$ implies that (x^k, y^*) is feasible for (1.2) for all k sufficiently large. Consequently we have $f(x^k) > f(x^*)$, due to $x^k \neq x^*$. By part (a), there exist $r > 0$, $T > 0$ such that for all $t \in (0, T]$ the regularised problem NLP(t) has a local minimum (x^t, y^t) with $x^t \in B_r(x^*)$. □

If (x^*, y^*) is a strict local minimum of the reformulation (1.2) with respect to x, but the cardinality constraint is not active, then Theorem 4.27 does not guarantee the existence of solutions of NLP(t) in a neighbourhood unless x^* is a strict local minimum of the original problem (1.1). This is in fact an advantage because local minima of the reformulation (1.2), in which the cardinality constraint is not active, are not necessarily local minima of the original problem (1.1). Hence those are no points we want the regularisation method to converge to. Precisely this situation is illustrated in the following example.

Example 4.28. Consider the cardinality constrained optimization problem

$$\min_{x \in \mathbb{R}^3} f(x) = \|x - (0, 1, 2)^T\|^2 \quad \text{s.t.} \quad \|x\|_0 \leq 1.$$

Then $x^1 = (0, 0, 2)^T$ is the global minimum, $x^2 = (0, 1, 0)^T$ is a local minimum, but $x^* = (0, 0, 0)^T$ is no local minimum. Now consider the continuous reformulation, which is is given by

$$\min_{x \in \mathbb{R}^3, \, y \in \mathbb{R}^3} \; f(x) = \|x - (0, 1, 2)^T\|^2 \quad \text{s.t.} \quad 0 \leq y \leq e, \quad e^T y \geq 2, \quad x \circ y = 0.$$

Then, choosing $y^* = (1, 1, 1)^T$, the point (x^*, y^*) is a strict local minimum of the continuous reformulation with respect to x since for all $r \in (0, 1)$ all points $(x, y) \in B_r(x^*, y^*) \cap Z$ satisfy $x = x^*$. The regularised program for a parameter $t > 0$ is given by

$$\min_{x \in \mathbb{R}^3, \, y \in \mathbb{R}^3} \; f(x) = \|x - (0, 1, 2)^T\|^2 \quad \text{s.t.} \quad 0 \leq y \leq e, \quad e^T y \geq 2, \quad -te \leq x \circ y \leq te.$$

For all $(x, y) \in Z(t)$ sufficiently close to (x^*, y^*) we have $y_i > 0$ and $e^T y > 2$. Hence in a neighbourhood of (x^*, y^*) the KKT conditions of the regularised program in (x, y) imply

$$\begin{aligned} 0 &= 2(x_2 - 1) + \gamma_2 y_2 \quad \Longrightarrow \quad \gamma_2 y_2 \approx 2, \\ 0 &= 2(x_3 - 2) + \gamma_3 y_3 \quad \Longrightarrow \quad \gamma_3 y_3 \approx 4, \\ 0 &= \nu + \gamma \circ x, \quad \nu \geq 0, \quad \gamma \circ x \geq 0. \end{aligned}$$

Here, the last equation implies $\nu = 0$ and $\gamma \circ x = 0$, which is only possible if $\gamma = 0$. This, however, is a contradiction to the first two equations. Thus the KKT conditions cannot be satisfied in a neighbourhood of (x^*, y^*). Since CC-LICQ holds in (x^*, y^*), it follows from Theorem 4.26 that MFCQ holds for the regularised problem sufficiently close to (x^*, y^*). Consequently the regularised program cannot have local minima in a vicinity of (x^*, y^*).
This implies that the Scholtes-type regularisation cannot converge to the undesirable local solution (x^*, y^*) of the continuous reformulation, which does not correspond to a solution of the original problem.

We are now in a position to prove the main result for this regularisation: Whenever x^* is a strict local minimum of (1.1) satisfying CC-MFCQ, then the Scholtes-type regularisation method is locally well defined. Furthermore, as $t^k \downarrow 0$, the KKT points (x^k, y^k) of NLP(t^k) converge to x^* at least in the x-component. If additionally $\|x^*\|_0 = \kappa$ holds, then the y-component is also convergent. To prove the result we can combine the previous results from this section and second order conditions from Section 3.3.

Theorem 4.29. *(a) Let x^* be a strict local minimiser of* (1.1) *(or (x^*, y^*) be a strict local minimum of* (1.2) *with respect to x and $\|x^*\|_0 = \kappa$) such that CC-MFCQ holds in x^*. Then there exist $T > 0$ and $r > 0$ such that for all $t \in (0, T]$ the regularised problem NLP(t) has a local minimum/KKT point (x^t, y^t) with $x^t \in B_r(x^*)$.*

(b) Let $(x^*, y^*) \in Z$ satisfy CC-MFCQ and choose $r > 0$ sufficiently small. Consider a sequence $(t_k)_k \downarrow 0$ and KKT points $(x^k, y^k)_k$ of NLP(t_k) such that $x^k \in B_r(x^*)$ for all $k \in \mathbb{N}$. Then the sequence $(x^k, y^k)_k$ has accumulation points and every accumulation point $(\bar{x}, \bar{y})$ is an S-stationary point of (1.2)*.*

(c) Let f, g, h be twice continuously differentiable. Let (x^*, y^*) be a strict local minimum of (1.2) *with respect to x. Let further (x^*, y^*) satisfy CC-MFCQ, $\|x^*\|_0 = \kappa$, and*

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x^*) \right) d_x > 0$$

hold for all $(d_x, d_y) \in \mathcal{C}_Z^{CC}(x^*, y^*)$ with $d_x \neq 0$ and all S-stationary multipliers (λ, μ, γ) of (x^*, y^*). Then there exists $r > 0$ such that for all sequences $(t_k)_k \downarrow 0$ and for all k sufficiently large, the problem NLP(t_k) has a KKT point (x^k, y^k) with $x^k \in B_r(x^*)$ and $(x^k, y^k) \to (x^*, y^*)$.

(d) Let f, g, h be twice continuously differentiable. Let x^* be a strict local minimum of (1.1) *such that CC-MFCQ holds and*

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x^*) \right) d_x > 0$$

hold for all $d_x \in \mathcal{C}_{\mathcal{X}}(x^*)$ with $d_x \neq 0$ and all M-stationary multipliers (λ, μ, γ) of x^*. Then there exists $r > 0$ such that for all sequences $(t_k)_k \downarrow 0$ for all k sufficiently large NLP(t_k) has a KKT point (x^k, y^k) with $x^k \in B_r(x^*)$ and $x^k \to x^*$.

Proof. (a) By Theorem 4.27 the assumptions guarantee the existence of $T > 0$ and $r > 0$ such that for all $t \in (0, T]$ the regularised problem NLP(t) has a local minimum (x^t, y^t) with $x^t \in B_r(x^*)$. Decreasing T and r if necessary we can also use Theorem 4.26, which guarantees MFCQ for NLP(t) in (x^t, y^t) and thus ensures that (x^t, y^t) are KKT points.

(b) Since $x^k \in B_r(x^*)$ and $y^k \in [0,1]^n$ for all $k \in \mathbb{N}$, the sequence $(x^k, y^k)_k$ is bounded and thus has at least one accumulation point. Now consider an arbitrary accumulation point $(\bar{x}, \bar{y})$. Since $t_k \downarrow 0$ we know that $(\bar{x}, \bar{y})$ is feasible for (1.2). If we chose $r > 0$ small enough, Lemma 3.28 tells us that CC-MFCQ in x^* implies CC-MFCQ in $(\bar{x}, \bar{y})$. Thus by Theorem 4.25 $(\bar{x}, \bar{y})$ is an S-stationary point of (1.2).

(c) Combining part (a) and (b), we see that there exists $r > 0$ such that for all $t_k > 0$ sufficiently small NLP(t_k) has a KKT point (x^k, y^k) with $x^k \in B_r(x^*)$ and that all accumulation points $(\bar{x}, \bar{y})$ of $(x^k, y^k)_k$ are S-stationary and thus M-stationary. By choosing $r > 0$ small enough, we can enforce $\bar{y} = y^*$. Since (x^*, y^*) is a local minimum satisfying CC-MFCQ, it is an S-stationary point, too. Furthermore, since $\|x^*\|_0 = \kappa$, S- and M-stationarity coincide and thus (x^*, y^*) satisfies the assumptions for Theorem 3.50. Thus, if we choose $r > 0$ small M-stationarity of the accumulation points $(\bar{x}, \bar{y}) = (\bar{x}, y^*)$ implies $\bar{x} = x^*$. This shows $x^k \to x^*$ and $y^k \to y^*$.

(d) Since x^* is a local minimum of (1.1), every $(x^*, y) \in Z$ is a local minimum of (1.2) and, due to CC-MFCQ, an S-stationary and thus M-stationary point. Furthermore, the set of M-stationary multipliers is independent from y. Using the assumptions, we obtain from Corollary 3.51 that there exists an $r > 0$ such that all M-stationary points $(\bar{x}, \bar{y}) \in Z$ with $\bar{x} \in B_r(x^*)$ satisfy $\bar{x} = x^*$. Analogously to (c) we see that we can decrease $r > 0$ such that for all $t_k > 0$ sufficiently small NLP(t_k) has a KKT point (x^k, y^k) with $x^k \in B_r(x^*)$ and that all accumulation points $(\bar{x}, \bar{y})$ of $(x^k, y^k)_k$ are S-stationary and thus M-stationary. Consequently all accumulation points satisfy $\bar{x} = x^*$ which shows $x^k \to x^*$.

$\square$

Note that the second order condition in part (c) and (d) is automatically satisfied if f is uniformly convex, g convex and h affine linear. Furthermore, in part (c) the additional assumption $\|x^*\|_0 = \kappa$ implies by Theorem 2.4 that the vector x^* is a strict local minimum of the cardinality constrained problem (1.1). Combining this with a few other previously used arguments, one can alternatively prove part (c) by showing that it is implied by part (d).

Remark 4.30. An inspection of the proof of Theorem 4.27 reveals that it is in fact independent from the particular constraints

$$-t \leq x_i \cdot y_i \leq t, \quad i = 1, \dots, n,$$

which replace the complementarity constraint for the Scholtes-type regularisation. It uses merely the fact that the limit of a sequence of feasible points (for decreasing regularisation

parameters) of the regularised problems, is feasible for (1.2). This means Theorem 4.27 holds for every regularisation method with said property.
Assuming that a suitable CC-constraint qualification holds, Theorem 4.29 can thus be adapted to any regularisation method for which two results hold: Firstly, the method should be convergent to M- or S-stationary points, i.e. a result corresponding to Theorem 4.25 is required. Secondly, a constraint qualification should hold for the regularised problems locally around said stationary point. This means a result corresponding to Theorem 4.26 is required as well.
In the following section we consider a different regularisation method which fulfils these requirements.

An important follow up question to the convergence of a regularisation method is the convergence in the inexact case. In practice solvers applied to the regularised problems compute only an approximation of a KKT point. Whether the limit of such a sequence of approximate KKT points is still S- or at least M-stationary is an open question. In the MPCC case, the Scholtes regularisation retains its convergence properties even if one computes only approximate KKT points of the regularised problems. This is an advantage over other regularisation methods, some of which have even stronger convergence properties in the exact case, see [55]. Since the CC-constraint qualifications are only formulated with respect to the x-variable of the complementarity formulation, we cannot use the same line of argument, as in the proof for the MPCC result. Therefore, it remains an open question at this moment, whether the Scholtes-type regularisation possesses this advantage also regarding the application to the complementarity formulation.

4.3.2 Kanzow-Schwartz Regularisation

The second regularisation method we are considering was also originally studied for MPCCs [54]. It was adapted to the complementarity formulation of cardinality constrained optimization problems in [14]. To relax the complementarity constraint let

$$\varphi_i^-(x,y;t) := \begin{cases} (-x_i - t)\cdot(y_i - t), & \text{if } -x_i + y_i \geq 2t, \\ -\frac{1}{2}\left((-x_i - t)^2 + (y_i - t)^2\right), & \text{if } -x_i + y_i < 2t, \end{cases}$$
$$\varphi_i^+(x,y;t) := \begin{cases} (x_i - t)\cdot(y_i - t), & \text{if } x_i + y_i \geq 2t, \\ -\frac{1}{2}\left((x_i - t)^2 + (y_i - t)^2\right), & \text{if } x_i + y_i < 2t, \end{cases}$$

for $x, y \in \mathbb{R}^n$ and $t \geq 0$. Because in the MPCC case one only considers the nonnegative part of the abscissa, only the functions φ_i^+, $i = 1, \ldots, n$, are required. Since in (1.2) there is no nonnegativity constraint on the variable x, the functions φ_i^-, $i = 1, \ldots, n$, are added. It is easy to confirm that

$$\varphi_i^-(x,y;t) \leq 0 \quad \Leftrightarrow \quad \min\{-x_i, y_i\} \leq t \quad \text{and} \quad \varphi_i^+(x,y;t) \leq 0 \quad \Leftrightarrow \quad \min\{-x_i, y_i\} \leq t.$$

The relaxation is illustrated in Figure 4.5 for different regularisation parameters. Furthermore, by [14, Lemma 5.1.] the functions φ_i^-, φ_i^+, $i = 1, \ldots, n$, are continuously differentiable, which is essential for the convergence analysis. Replacing the complementarity constraint in (1.2) results in the regularised problem

$$\begin{aligned} \mathrm{NLP}(t): \quad \min_{x,y}\ f(x) \quad \text{s.t.} \quad & g_i(x) \le 0, && \forall i = 1, \dots, m, \\ & h_i(x) = 0, && \forall i = 1, \dots, p, \\ & \sum\nolimits_{i=1}^{n} y_i \ge n - \kappa, && \\ & \varphi_i^-(x, y; t) \le 0, && \forall i = 1, \dots, n, \\ & \varphi_i^+(x, y; t) \le 0, && \forall i = 1, \dots, n, \\ & 0 \le y_i \le 1, && \forall i = 1, \dots, n. \end{aligned} \tag{4.52}$$

As in the introduction to this section, we denote the feasible set of $\mathrm{NLP}(t)$ by $Z(t)$ for $t \ge 0$. We have $Z(t_2) \subseteq Z(t_1)$ for all $0 \le t_2 \le t_1$ as well as $Z \subseteq Z(t)$ for all $t \ge 0$. The following convergence result holds for the Kanzow-Schwartz regularisation.

Theorem 4.31 ([14, Theorem 5.3])**.** *Let $(t^k)_{k\in\mathbb{N}} \subseteq \mathbb{R}_+$, $t^k \downarrow 0$ $(k \to \infty)$ and $(x^k, y^k)_{k\in\mathbb{N}}$ be a sequence of KKT points of $\mathrm{NLP}(t^k)$ such that $(x^k, y^k) \to (x^*, y^*)$ as $k \to \infty$. Assume that CC-CPLD holds at (x^*, y^*). Then x^* as an M-stationary point of* (1.2)*.*

The above convergence result holds under the CC-CPLD constraint qualification, which is rather weak by comparison to the other CC-constraint qualifications (see Section 3.2.1). This result is similar to the convergence result for this regularisation for MPCCs (see [46, Theorem 3.13]) for which also a CPLD-type constraint qualification is required.

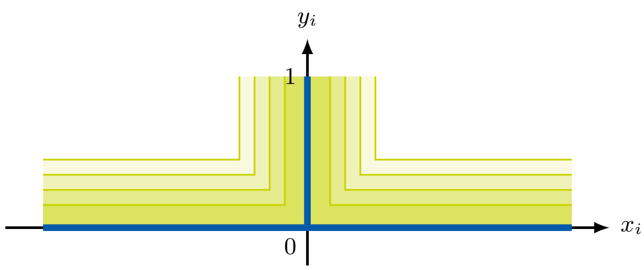


Figure 4.5: Pairs (x_i, y_i) that fulfil the complementarity constraint (blue) and the relaxed complementarity constraint (green) for the Kanzow-Schwartz regularisation. The relaxation is illustrated for decreasing regularisation parameters $t_1 > t_2 > \dots$ (darker green for smaller parameters).

Furthermore, the following result holds for this regularisation: If CC-CPLD is satisfied at a feasible point of (1.2), then for sufficiently small regularisation parameters GCQ holds for $\mathrm{NLP}(t)$ in a vicinity of this point.

Theorem 4.32 ([14, Theorem 5.4])**.** *Let (x^*, y^*) be feasible for* (1.2) *and satisfy CC-CPLD. Then there exist $T > 0$ and $r > 0$ such that for all $t \in (0, T]$ standard GCQ for NLP(t) holds at every $(x, y) \in Z(t)$ with $(x, y) \in B_r(x^*) \times B_r(y^*)$.*

In contrast to the above theorem, a considerably stronger LICQ-type constraint qualification is required for the corresponding result for this regularisation for MPCCs (see [46, Theorem 3.14]).

In the previous section we were able to extend the convergence theory of the Scholtes-type regularisation using second order conditions from Section 3.3. We will investigate next if a similar result can be obtained for the Kanzow-Schwartz regularisation. By Remark 4.30, Theorem 4.27 also holds for the Kanzow-Schwartz regularisation. Taking Theorem 4.31 and Theorem 4.32 into consideration, we thus can adapt Theorem 4.29 to the Kanzow-Schwartz regularisation.

Theorem 4.33. *(a) Let x^* be a strict local minimiser of* (1.1) *(or (x^*, y^*) be a strict local minimum of* (1.2) *with respect to x and $\|x^*\|_0 = \kappa$) such that CC-CPLD holds in x^*. Then there exist $T > 0$ and $r > 0$ such that for all $t \in (0, T]$ the regularised problem NLP(t) has a local minimum/KKT point (x^t, y^t) with $x^t \in B_r(x^*)$.*

(b) Let $(x^*, y^*) \in Z$ satisfy CC-CPLD and choose $r > 0$ sufficiently small. Consider a sequence $t_k \downarrow 0$ and KKT points $(x^k, y^k)_k$ of NLP(t_k) such that $x^k \in B_r(x^*)$ for all $k \in \mathbb{N}$. Then the sequence $(x^k, y^k)_k$ has accumulation points and every accumulation point $(\bar{x}, \bar{y})$ is an M-stationary point of (1.2).

(c) Let f, g, h be twice continuously differentiable. Let (x^*, y^*) be a strict local minimum of (1.2) *with respect to x and $\|x^*\|_0 = \kappa$ such that CC-CPLD holds and*

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x^*) \right) d_x > 0$$

hold for all $(d_x, d_y) \in \mathcal{C}_Z^{CC}(x^*, y^*)$ with $d_x \neq 0$ and all S-stationary multipliers (λ, μ, γ) of (x^*, y^*). Then there exists $r > 0$ such that for all sequences $t_k \downarrow 0$ for all k sufficiently large NLP(t_k) has a KKT point (x^k, y^k) with $x^k \in B_r(x^*)$ and $(x^k, y^k) \to (x^*, y^*)$.

(d) Let f, g, h be twice continuously differentiable. Let x^* be a strict local minimum of (1.1) *such that CC-CPLD holds and*

$$d_x^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x^*) \right) d_x > 0$$

hold for all $d_x \in \mathcal{C}_{\mathcal{X}}(x^*)$ with $d_x \neq 0$ and all M-stationary multipliers (λ, μ, γ) of x^*. Then there exists $r > 0$ such that for all sequences $t_k \downarrow 0$ for all k sufficiently large NLP(t_k) has a KKT point (x^k, y^k) with $x^k \in B_r(x^*)$ and $x^k \to x^*$.

The proof of Theorem 4.33 is analogue to the proof of Theorem 4.29. We don't need to presume CC-MFCQ, since Theorem 4.31 and Theorem 4.32 require only CC-CPLD. Consequently the accumulation points in part (b) are only M-stationary. For the proof of part (b) we use Lemma 3.29 to show that CC-CPLD holds in a neighbourhood of x^*. The proof of parts (c) and (d) is then analogue to the proof of Theorem 4.29 (c) and (d).
We will consider the Kanzow-Schwartz regularisation again in Chapter 5.

4.3.3 Exponential Regularisation

In this section we consider a regularisation that uses exponential functions to relax the complementarity constraints. The results on this regularisation were, independently from this thesis, also derived in [42], where a more detailed study of this method can be found.

For a regularisation parameter $t > 0$ they are replaced by

$$y_i \le e^{-t \cdot x_i} \quad \text{and} \quad y_i \le e^{t \cdot x_i} \tag{4.53}$$

for $i = 1, \ldots, n$, see Figure 4.6 for an illustration for different regularisation parameters. Here, we increase the regularisation parameter t to better approximate the cardinality constraints. The result of the relaxation with the above exponential functions is the regularised problem

$$\begin{aligned} \text{NLP}(t): \quad \min_{x,y} \ f(x) \quad \text{s.t.} \quad & g_i(x) \le 0, && \forall i = 1, \ldots, m, \\ & h_i(x) = 0, && \forall i = 1, \ldots, p, \\ & \sum\nolimits_{i=1}^{n} y_i \ge n - \kappa, \\ & y_i \le e^{-t \cdot x_i}, && \forall i = 1, \ldots, n, \\ & y_i \le e^{t \cdot x_i}, && \forall i = 1, \ldots, n, \\ & 0 \le y_i \le 1, && \forall i = 1, \ldots, n. \end{aligned} \tag{4.54}$$

As before, we denote the feasible set of NLP(t) by $Z(t)$. We have $Z \subseteq Z(t)$ for all $t > 0$ and

$$Z(t_2) \subseteq Z(t_1) \quad \forall\ 0 < t_1 \le t_2.$$

Thus, unlike as for the Scholtes-type regularisation and the Kanzow-Schwartz regularisation, the approach here will be to compute a sequence of KKT points of NLP(t) for *increasing* parameters t. Figure 4.6 highlights a further difference to the previous two regularisation approaches: The relaxing constraints (4.53) can be active in the same point. But this is the case if and only if $y_i = 1$ in (4.53) for a given $i \in \{1, \ldots, n\}$. We will use this fact in the proof of Theorem 4.35.

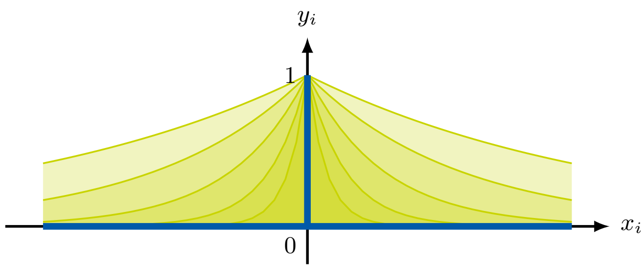


Figure 4.6: Pairs (x_i, y_i) that fulfil the complementarity constraint (blue) and the relaxed complementarity constraint (green) for the exponential regularisation. The relaxation is illustrated for increasing regularisation parameters $0 < t_1 < t_2 < \ldots$ (darker green for larger parameters).

In [1] this type of relaxation is used in the context of chance constrained optimization problems. The approach therein is to reformulate the problem in a similar fashion to (1.1) using a complementarity formulation. It then is argued that the application of standard constraint qualifications is not sensible. A regularisation method that relaxes the complementarity constraint in a similar way to (4.53) is studied. The convergence to a point which fulfils a first order optimality condition is shown under a constraint qualification similar to CC-MFCQ [1, Theorem 2.2, Theorem 3.1]. For the regularisation of (1.2) we can derive the following result, see also [42].

Theorem 4.34. *Let $(t_k)_{k\in\mathbb{N}}$ be a sequence with $t_k > 0$ for all $k \in \mathbb{N}$ and $t_k \to \infty$. Let $(x^k, y^k)_{k\in\mathbb{N}}$ be a sequence of KKT points of* $\mathrm{NLP}(t_k)$ *converging to (x^*, y^*). If CC-MFCQ holds in (x^*, y^*), then (x^*, y^*) is an S-stationary point of* (1.2).

Proof. Since $(x^k, y^k)_{k\in\mathbb{N}}$ is a sequence of KKT points, for all $k \in \mathbb{N}$ there are multipliers $(\lambda^k, \mu^k, \gamma^{k-}, \gamma^{k+}, \delta^k, \nu^k)$ such that we have

$$\nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla h_i(x^k) + \sum_{i=1}^{n} (\gamma_i^{k+} - \gamma_i^{k-}) t_k e^{-t_k|x_i^k|} e_i = 0, \tag{4.55}$$

$$-\delta^k e - \sum_{i=1}^{n} \nu_i^k e_i + \sum_{i=1}^{n} (\gamma_i^{k+} + \gamma_i^{k-}) e_i = 0, \tag{4.56}$$

$$\begin{aligned} \lambda_i^k \geq 0,\ \lambda_i^k \cdot g_i(x^k) = 0, &\quad \forall i = 1, \dots, m,\\ \gamma_i^{k-} \geq 0,\ \gamma_i^{k-} \cdot (y_i^k - e^{t_k x_i^k}), &\quad \forall i = 1, \dots, n,\\ \gamma_i^{k+} \geq 0,\ \gamma_i^{k+} \cdot (y_i^k - e^{-t_k x_i^k}), &\quad \forall i = 1, \dots, n,\\ \delta^k \geq 0,\ \delta^k \cdot (e^T y^k - n + \kappa) = 0, &\\ \nu_i^k \geq 0,\ \nu_i^k \cdot y_i^k = 0, &\quad \forall i = 1, \dots, n. \end{aligned}$$

Assume there are indexes $i \in \{1, \dots, n\}$ and $k \in \mathbb{N}$ such that $\nu_i^k > 0$. Then $y_i^k = 0$, thus $\gamma_i^{k-} = \gamma_i^{k+} = 0$ and from (4.56) we have $0 > -\nu_i^k = \delta^k \geq 0$, which is a contradiction. Therefore $\nu^k = 0$ for all $k \in \mathbb{N}$ and (4.56) simplifies to

$$\delta^k = \gamma_i^{k-} + \gamma_i^{k+} \quad \forall i = 1, \dots, n.$$

Thus if $\delta^k = 0$, we have $\gamma_i^{k-} = \gamma_i^{k+} = 0$ for all $i = 1, \dots, n$. If $\delta^k > 0$ we have $y_i^k = e^{-t_k|x_i^k|}$ for all $i = 1, \dots, n$.
Let $\gamma_i^k := (\gamma_i^{k+} - \gamma_i^{k-}) t_k e^{-t_k|x_i^k|}$ for all $i = 1, \dots, n$ and all $k \in \mathbb{N}$.
We show by contradiction that the sequence $(\lambda^k, \mu^k, \gamma^k)_{k\in\mathbb{N}}$ is bounded. To this end assume that $(\lambda^k, \mu^k, \gamma^k)_k$ is unbounded. Then the sequence

$$\left(\frac{(\lambda^k, \mu^k, \gamma^k)}{\|(\lambda^k, \mu^k, \gamma^k)\|} \right)_{k\in\mathbb{N}}$$

is bounded (if necessary, we consider a subsequence with $(\lambda^k, \mu^k, \gamma^k) \neq 0$). Without loss of generality let it be convergent and

$$0 \neq \lim_{k\to\infty} \frac{(\lambda^k, \mu^k, \gamma^k)}{\|(\lambda^k, \mu^k, \gamma^k)\|} =: (\bar{\lambda}, \bar{\mu}, \bar{\gamma}).$$

Because $\lambda_i^k \geq 0$ for all $i = 1, \dots, m$ and all $k \in \mathbb{N}$, we have $\bar{\lambda} \geq 0$. If $g_i(x^*) < 0$ holds, then $g_i(x^k) < 0$, hence $\lambda_i^k = 0$, for sufficiently large k, and thus $\bar{\lambda}_i = 0$. Thus we also have $\mathrm{supp}(\bar{\lambda}) \subseteq I_g(x^*)$.
We will show $\mathrm{supp}(\bar{\gamma}) \subseteq I_0(x^*)$ by contradiction. Assume there is an index $j \in \{1, \dots, n\}$ such that $\bar{\gamma}_j \neq 0$ and $x_j^* \neq 0$. Because (x^*, y^*) is feasible for (1.2), we have $y_j^* = 0$. Since $\bar{\gamma}_j \neq 0$, we have

$$\frac{(\gamma_j^{k+} - \gamma_j^{k-}) t_k e^{-t_k|x_j^k|}}{\|(\lambda^k, \mu^k, \gamma^k)\|} \neq 0$$

and thus $\gamma_j^{k+} - \gamma_j^{k-} \neq 0$ for sufficiently large k. This implies $\delta^k > 0$. Taking into account our previous considerations, we consequently have $e^T y^k = n - \kappa$, and

$$\begin{aligned} 0 < \delta^k = \gamma_i^{k+} - \gamma_i^{k-}, &\quad \forall i = 1, \dots, n, \\ y_j^k = t_k e^{-t_k |x_j^k|}, &\quad \forall i = 1, \dots, n, \end{aligned}$$

for sufficiently large k. Particularly this means $y_j^k > 0$ for all k sufficiently large. Since $y_j^k \to y_j^* = 0$ $(k \to \infty)$, while $e^T y^k = n - \kappa$ for sufficiently large k, there is an index $l \in \{1, \dots, n\} \setminus \{j\}$ such that we have (at least for a subsequence)

$$y_l^k < y_l^{k+1}.$$

Then we have $y_l^k \in (0, 1)$ for all k sufficiently large and $y_l^* > 0$. Because $y_i^k < 1$, $i = j, l$, only one of the constrains $y_i \leq e^{t_k x_i^k}$ or $y_i \leq e^{-t_k x_i^k}$ is active. Thus

$$|\gamma_i^{k+} - \gamma_i^{k-}| = \gamma_i^{k+} + \gamma_i^{k-} = \delta^k \quad i = j, l.$$

Using the above equality, we can deduce

$$\frac{|\bar{\gamma}_j|}{|\bar{\gamma}_l|} = \lim_{k\to\infty} \frac{\gamma_j^k}{\gamma_l^k} = \lim_{k\to\infty} \frac{|\gamma_j^{k+} - \gamma_j^{k-}| t_k e^{-t_k |x_j^k|}}{|\gamma_l^{k+} - \gamma_l^{k-}| t_k e^{-t_k |x_l^k|}} = \lim_{k\to\infty} \frac{\delta^k y_j^k}{\delta^k y_l^k} = \lim_{k\to\infty} \frac{y_j^k}{y_l^k} = \frac{y_j^*}{y_l^*} = 0.$$

Because we assumed $\bar{\gamma}_j \neq 0$, this is a contradiction. We consequently have

$$\mathrm{supp}(\bar{\gamma}) \subseteq I_0(x^*).$$

From dividing (4.55) by $\|\lambda^k, \mu^k, \gamma^k\|$, taking the limit $k \to \infty$ and paying regard to $\mathrm{supp}(\bar{\lambda}) \subset I_g(x^*)$ and $\mathrm{supp}(\bar{\gamma}) \subseteq I_0(x^*)$, we obtain

$$\sum_{i \in I_g(x^*)} \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^{p} \bar{\mu}_i \nabla h_i(x^*) + \sum_{i \in I_0(x^*)} \bar{\gamma}_i e_i = 0.$$

Since $\bar{\lambda} \geq 0$ and $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}) \neq 0$, this is a contradiction to CC-MFCQ in (x^*, y^*). Hence our initial assumption was wrong and the sequence $(\lambda^k, \mu^k, \gamma^k)_{k\in\mathbb{N}}$ is bounded. Without loss of generality let $(\lambda^k, \mu^k, \gamma^k)_{k\in\mathbb{N}}$ be convergent and

$$\lim_{k\to\infty} (\lambda^k, \mu^k, \gamma^k) =: (\lambda, \mu, \gamma).$$

Taking the limit $k \to \infty$ in (4.55) yields

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i=1}^{n} \gamma_i e_i = 0. \tag{4.57}$$

As before, we obtain $\lambda \geq 0$ and $\mathrm{supp}(\lambda) \subseteq I_g(x^*)$.
Lastly we will show $\mathrm{supp}(\gamma) \subseteq I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$. Let $j \in I_{00}(x^*, y^*) \cup I_{\pm 0}(x^*, y^*)$. In case we have $y_j^k = 0$ for all k sufficiently large, it follows that $\gamma_j^{k-} = \gamma_j^{k+} = 0$ and thus $\gamma_j = 0$.

Otherwise we have $y_j^k \to 0+$ $(k \to \infty)$. Then either we have $\delta^k = 0$, hence $\gamma_j^{k-} = \gamma_j^{k+} = 0$ for large k and thus $\gamma_j = 0$. Or we have $\delta^k > 0$ and hence $e^T y^k = n - \kappa$ for all k sufficiently large. As before, this implies the existence of an index l such that $y_l^k \to y_l^* \in (0, 1]$ and

$$\lim_{k\to\infty} \frac{\gamma_j^k}{\gamma_l^k} = 0,$$

thus also $\gamma_j = 0$. We thus have $\operatorname{supp}(\gamma) \subseteq I_{0+}(x^*, y^*) \cup I_{01}(x^*, y^*)$ and by (4.57) the limit (x^*, y^*) is S-stationary. □

We continue with a result on the validity of constraint qualifications for the regularised problems. Similar to the Scholtes-type regularisation, MFCQ holds for NLP(t) in the vicinity of a point fulfilling CC-MFCQ. For the following result, compare also [42].

Theorem 4.35. *Let $t > 0$ and CC-MFCQ hold in (x^*, y^*). Then there is a $T > 0$ and a neighbourhood $N(x^*, y^*)$ of (x^*, y^*) such that for all $t \in (0, T]$ MFCQ holds in every $(x, y) \in N(x^*, y^*) \cap Z(t)$ for NLP(t).*

Proof. We assume that the claim is wrong. Then we can find sequences $(t_k)_{k\in\mathbb{N}}$, $(x^k, y^k)_{k\in\mathbb{N}}$ and $(\lambda^k, \mu^k, \delta^k, \nu^k, \gamma^{-k}, \gamma^{+k})_{k\in\mathbb{N}}$ such that for all $k \in \mathbb{N}$ the point (x^k, y^k) is feasible for NLP(t_k) and we have $t_k > 0$, $t_k \to \infty$ $(k \to \infty)$, $(\lambda^k, \mu^k, \delta^k, \nu^k, \gamma^{-k}, \gamma^{+k}) \neq 0$, as well as

$$\sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) + \sum_{i=1}^n (\gamma_i^{k+} - \gamma_i^{k-}) t_k e^{-t_k |x_i^k|} e_i = 0, \tag{4.58}$$

$$-\delta^k e - \sum_{i=1}^n \nu_i^k e_i + \sum_{i=1}^n (\gamma_i^{k+} + \gamma_i^{k-}) e_i = 0, \tag{4.59}$$

$$\begin{aligned} \lambda_i^k \geq 0,\ \lambda_i^k \cdot g_i(x^k) = 0, &\quad \forall i = 1, \dots, m, \\ \gamma_i^{k-} \geq 0,\ \gamma_i^{k-} \cdot (y_i^k - e^{t_k x_i^k}), &\quad \forall i = 1, \dots, n, \\ \gamma_i^{k+} \geq 0,\ \gamma_i^{k+} \cdot (y_i^k - e^{-t_k x_i^k}), &\quad \forall i = 1, \dots, n, \\ \delta^k \geq 0,\ \delta^k \cdot (e^T y^k - n + \kappa) = 0, & \\ \nu_i^k \geq 0,\ \nu_i^k \cdot y_i^k = 0, &\quad \forall i = 1, \dots, n. \end{aligned}$$

This means the relevant gradients are positive linearly dependent in (x^k, y^k), thus MFCQ is violated. We will deduce a contradiction to CC-MFCQ in (x^*, y^*) from this.
Equation (4.59) can be simplified. Assume there are indexes $i \in \{1, \dots, n\}$ and $k \in \mathbb{N}$ such that $\nu_i^k > 0$. Then we have $y_i^k = 0$ and thus $\gamma_i^{k-} = \gamma_i^{k+} = 0$. From (4.59) we thus obtain $0 \leq \delta^k = -\nu_i^k < 0$, a contradiction. We consequently have

$$\nu^k = 0 \quad \forall k \in \mathbb{N}$$

and can write (4.59) as

$$\delta^k = \gamma_i^{k+} - \gamma_i^{k-} \quad \forall i = 1, \dots, n. \tag{4.60}$$

Let $\gamma_i := (\gamma_i^{k+} - \gamma_i^{k-}) t_k e^{-t_k |x_i^k|}$ for all $i = 1, \dots, n$.
To show for all $i = 1, \dots, n$ that $\gamma_i \neq 0$ holds, if $\gamma_i^{k-} \neq 0$ or $\gamma_i^{k+} \neq 0$, we distinguish the following cases.

Firstly, if $\gamma_i^{k-} > 0 = \gamma_i^{k+}$, $\gamma_i^{k-} = 0 < \gamma_i^{k+}$, or $0 < \gamma_i^{k-} \neq \gamma_i^{k-} > 0$, then we have $\gamma_i^{k+} - \gamma_i^{k-} \neq 0$ thus $\gamma_i \neq 0$.
For the second case, assume $\gamma_j^{k-} = \gamma_j^{k+} > 0$ for an index $j \in \{1, \dots, n\}$. Then we have $e^{-t_k x_j^k} = y_j^k = e^{t_k x_j^k}$, thus $x_j^k = 0$ and $y_j^k = 1$. We also have $\delta^k = \gamma_j^{k-} + \gamma_j^{k+} > 0$, hence $e^T y^k = n - \kappa$ and thus

$$\left(y_i^k = e^{-t_k x_i^k} \quad \text{or} \quad y_i^k = e^{t_k x_i^k} \right) \quad \forall i = 1, \dots, n.$$

This means for all i at least one of the above constraints has to be active.
Consider the case that for some $l \in \{1, \dots, n\} \setminus \{j\}$ only one constraint is active. Without loss of generality let $y_l^k < e^{-t_k x_l^k}$ and $y_l^k = e^{t_k x_l^k}$. Then $\gamma_l^{k+} = 0$ and $0 < \delta^k = \gamma_l^{k-}$. We thus have $\gamma_l \neq 0$, hence $\gamma \neq 0$.
Now consider the case that $e^{-t_k x_i^k} = y_i^k = e^{t_k x_i^k}$ for all $i = 1, \dots, n$. Then $y_i^k = 1$ for all $i = 1, \dots, n$ and thus $e^T y^k = n > n - \kappa$, thus $\delta^k = 0$. This is a contradiction to $\delta^k = \gamma_j^{k-} + \gamma_j^{k+} \neq 0$.
Taking into account $\nu^k = 0$, we consequently have for all $k \in \mathbb{N}$ that

$$(\lambda^k, \mu^k, \delta^k, \nu^k, \gamma^{k-}, \gamma^{k+}) \neq 0 \quad \Rightarrow \quad (\lambda^k, \mu^k, \gamma^k, \delta^k) \neq 0.$$

Since $(\lambda^k, \mu^k, \gamma^k, \delta^k) \neq 0$ for all $k \in \mathbb{N}$, we can assume without loss of generality that $\|(\lambda^k, \mu^k, \gamma^k, \delta^k)\| = 1$. Without loss of generality let the sequence be convergent and

$$\lim_{k \to \infty} (\lambda^k, \mu^k, \gamma^k, \delta^k) =: (\lambda, \mu, \gamma, \delta) \neq 0.$$

If $\delta = 0$, we have $(\lambda, \mu, \gamma) \neq 0$.
If $\delta > 0$, then $\delta^k > 0$ for all k sufficiently large. From (4.60) we have $0 < \delta^k = \gamma_i^{k+} + \gamma_i^{k-}$ for all $i = 1, \dots, n$ and all k sufficiently large. From our previous considerations we have $\gamma^k \neq 0$ for all k sufficiently large. Thus $\gamma \neq 0$ and consequently, also in this case, we have

$$(\lambda, \mu, \gamma) \neq 0.$$

Since $\lambda_i^k \geq 0$ for all $i = 1, \dots, m$ and all $k \in \mathbb{N}$, we have $\lambda \geq 0$. If $g_i(x^*) < 0$ holds, then $g_i(x^k) < 0$, hence $\lambda_i^k = 0$, for sufficiently large k, and thus $\lambda_i = 0$. Hence we also have

$$\operatorname{supp}(\lambda) \subseteq I_g(x^*).$$

We will now show $\operatorname{supp}(\gamma) \subseteq I_0(x^*)$ by contradiction. To this end assume that there is an index $j \in \{1, \dots, n\}$ such that $\gamma_j \neq 0$ and $x_j^* \neq 0$. Then for all k sufficiently large we have $x_j^k \neq 0$ and $\gamma_j^k \neq 0$, hence $\gamma_j^{k+} - \gamma_j^{k-} \neq 0$. From (4.60) we then have $\delta^k > 0$ and hence $e^T y^k = n - \kappa$ for all k sufficiently large. In fact, this means we have $\gamma_i^{k+} + \gamma_i^{k-} = \delta^k > 0$ for all $i = 1, \dots, n$. Consequently, for k sufficiently large we have

$$\left(y_i^k = e^{-t_k x_i^k} \quad \text{or} \quad y_i^k = e^{-t_k x_i^k} \right) \quad \forall i = 1, \dots, n.$$

Thus $y_i^k > 0$ for all $i = 1, \dots, n$ and all k sufficiently large. Since (x^*, y^*) is feasible for (1.2), we have $y_j^* = 0$. Consequently $y_j^k \to 0+$ $(k \to \infty)$. Because we also have $e^T y^k = n - \kappa$ for k sufficiently large, there must exist another index $l \in \{1, \dots, n\}$ such that y_l^k is increasing, i.e.

$y_l^k < y_l^{k+1}$ (taking a subsequence of $(x^k, y^k)_k$ if necessary). This implies $y_l^k \in (0,1)$ for all k sufficiently large and $y_l^* = 1$.
Because for k sufficiently large we have $y_i^k < 1$ for $i = j, l$, not both constraints $y_i^k \leq e^{\pm t_k x_i^k}$ are active at the same time (for the indexes j and l). This means we have

$$|\gamma_i^{k+} - \gamma_i^{k-}| = |\gamma_i^{k+} + \gamma_i^{k-}| = \delta^k \quad \text{for } i = j, l,$$

and consequently

$$\frac{\gamma_j^k}{\gamma_l^k} = \frac{|\gamma_j^{k+} - \gamma_j^{k-}| t_k e^{-t_k |x_j^k|}}{|\gamma_l^{k+} - \gamma_l^{k-}| t_k e^{-t_k |x_l^k|}} = \frac{|\gamma_j^{k+} + \gamma_j^{k-}| y_j^k}{|\gamma_l^{k+} + \gamma_l^{k-}| y_l^k} = \frac{\delta^k y_j^k}{\delta^k y_l^k} = \frac{y_j^k}{y_l^k} \longrightarrow \frac{y_j^*}{y_l^*} = 0 \quad (k \to \infty).$$

This is a contradiction, since $\gamma_j^k \to \gamma_j \neq 0$ $(k \to \infty)$ by assumption. Thus we have

$$\text{supp}(\gamma) \subseteq I_0(x^*).$$

Using our definition of γ^k as well as $\text{supp}(\lambda) \subseteq I_g(x^*)$ and $\text{supp}(\gamma) \subseteq I_0(x^*)$, for $k \to \infty$ equation (4.58) yields

$$\sum_{i \in I_g(x^*)} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i \in I_0(x^*)} \gamma_i e_i = 0.$$

Because $\lambda \geq 0$ and $(\lambda, \mu, \gamma) \neq 0$, this is a contradiction to CC-MFCQ. □

Like for the previous regularisations, we can expand the the convergence theory of the exponential regularisations using the second order optimality conditions from Chapter 3. A Result corresponding to Theorem 4.29 holds for the exponential regularisation as well, see Remark 4.30. We will consider the exponential regularisation again in Chapter 5 for a numerical comparison of different methods for cardinality constrained optimization problems.

Concluding Remarks on the Numerical Methods

In this chapter we studied three types of numerical methods for the complementarity formulation that make use of the custom optimality conditions.
Firstly, we considered penalty methods: Using our result on the existence of a local error bound, we proved exactness of a distance-based penalty function. We showed that this result holds under different prerequisites: Assuming the functions f, ∇g and ∇h are locally Lipschitz continuous, the CC-CPLD constraint qualification is required. If the stronger CC-MFCQ holds, then the penalty function is exact without further assumptions on ∇g and ∇h.
Secondly, we considered an ℓ^1-penalty term for the complementarity constraint, in case the additional constraint $x \geq 0$ is present. Combined, the two main results for this approach, Theorem 4.14 and Theorem 4.16, give insight into the interplay of stationary conditions for the penalised problem and the complementarity formulation: If a KKT point of the penalised problem is feasible for the complementarity formulation, it is an S-stationary point. Moreover, if a sequence of KKT points of the penalised problem is convergent and its limit is feasible, we also know that it is S-stationary.
In Section 4.2 we considered a piecewise SQP scheme for the cardinality constrained optimization problem. An application of an SQP method to the complementarity formulation yields

	Prerequisite	CQ for auxiliary NLP	limit of KKT points
Penalty Methods			
ℓ^1-norm	CC-MFCQ	MFCQ	S-stationary, if feasible
distance-based	CC-CPLD and ∇g, ∇h loc. Lipschitz <u>or</u> CC-MFCQ	*exact penalty function*	
Regularisation Methods			
Scholtes	CC-MFCQ	MFCQ	S-stationary
Kanzow-Schwartz	CC-CPLD	GCQ	M-stationary
Exponential	CC-MFCQ	MFCQ	S-stationary

Table 4.1: Theoretical properties of the penalisation and regularisation methods.

quadratic subproblems which correspond to quadratic subproblems of a certain decomposition of the feasible set of (1.1). Using this decomposition, we then investigated the behaviour of a (standard) SQP method applied to the complementarity formulation.

Finally, we studied regularisation methods. We derived convergence results for the Scholtes-type regularisation and the exponential regularisation showing that the limit of a sequence of KKT points of regularised problems is S-stationary under CC-MFCQ. The result for the Scholtes-regularisation is stronger than the corresponding result for MPCCs: For MPCCs the limit only is known to be C-stationary under a MFCQ-type constraint qualification, which corresponds to M-stationary points in our setting, see Section 3.2.2. If one can show convergence to an S-stationary point for the Kanzow-Schwartz regularisation, assuming a stronger condition than CC-CPLD holds, remains an open question. Moreover, we used second order optimality conditions from the previous chapter to improve the convergence theory of the Scholtes-type regularisation. In fact, the procedure can be used to improve a whole class of regularisation methods, which we explicitly did for the Kanzow-Schwartz regularisation.

Table 4.1 summarises the theoretical properties of the penalty and regularisation methods.

5 Computational Results

In this chapter we present and discuss numerical results for the application of the complementarity formulation as a model for sparse portfolios. Furthermore, we present and discuss results of the numerical methods from Chapter 4 for a range of portfolio optimization problems.
In Section 5.1, we use the complementarity formulation for sparse portfolio selection. We use historical stock market data for the model and the Scholtes-type regularisation from Section 4.3.1 to compute a solution. We then evaluate the performance of the constructed portfolios by comparing their Sharpe ratio with the Sharpe ratio of an evenly distributed portfolio.
In Section 5.2 we test the penalty and regularisation methods from Chapter 4 on sparse portfolio optimization problems. For this comparison we use a set of randomly generated test problems and different types of risk measures as objective function. We also include a solver for (standard) nonlinear optimization problems in our study.

5.1 The Complementarity Formulation as a Model for Sparse Portfolios

In this section we use the complementarity formulation as a model for sparse portfolios and test its performance with historical stock market data. We consider portfolio optimization models which originate from the classic minimum-variance model introduced by Markowitz [61].
In this model an investor can choose from n possible assets. The vector $x \in \mathbb{R}^n$ represents the distribution of the investor's budget among these asset. The assets' returns are assumed to be randomly distributed. Let $r_t = (r_{1,t}, \dots, r_{n,t})^T \in \mathbb{R}^n$ be the returns at a given time t for each of the n assets. Let $\mu = \mathbb{E}[r_t]$ be the expected return and C the covariance matrix. A general form of the portfolio optimization problem is given by

$$\min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s.t.} \quad \begin{aligned} e^T x &= 1, \\ \mu^T x &= \rho. \end{aligned} \tag{5.1}$$

The investor can invest 100% of the budget, modelled by the first constraint. He or she aims to have an expected return of ρ, modelled by the second constraint, while minimising a risk measure $f \colon \mathbb{R}^n \to \mathbb{R}$. Further constraints, for instance lower and upper bounds can be added to the model. Inserting the variance $f(x) = x^T C x$ of the investments results in the basic minimum-variance model. The investor's strategy, i.e. the division of capital among the available assets, is given by a solution x^* of the optimization problem (5.1). Since the returns r_t assets are unknown they are estimated from historical data in practice. This exposes the above model and its solutions to estimation errors.
In [37], Giannone et al. propose the variance with an added ℓ^1-penalty term as objective function. Using the identity $C = \mathbb{E}[r_t^T r_t] - \mu^T \mu$ they reformulate the objective function (see

[37, Section 2]) and aim to solve the problem

$$\min_{x\in\mathbb{R}^n} \ \|\rho e_T - Rx\|_2^2 + \alpha\|x\|_1 \quad \text{s.t.} \quad \begin{aligned} e^T x &= 1, \\ \mu^T x &= \rho. \end{aligned} \tag{5.2}$$

The matrix $R \in \mathbb{R}^{T\times n}$ contains the returns of the assets $1,\dots,n$ over the period $1,\dots,T$, i.e. $R_{t,i} = (r_t)_i$. The vector $e_T \in \mathbb{R}^T$ consists of units and $\alpha > 0$ is a penalty parameter.

Giannone et al. argue that, due to the correlation of the data, the matrix R has a high condition number which leads to a high sensitivity to estimation errors. For inverse problems it is known that ℓ^1-penalisation has a regularising effect, see [20]. By adding the ℓ^1-penalty term to the above model, one expects to have a similar effect and thus to obtain better performing portfolios. The ℓ^1 term promotes sparsity of the solution: Portfolios with a small number of active positions are penalised less in the objective function. This can be desirable because the investor obtains a portfolio which is easier to manage, and can reduce transaction costs.

The performance of various strategies originating from the Markowitz model was investigated by DeMiguel et al. in [22]. The authors find that none of the tested models deliver a strategy that outperforms the heuristic strategy of evenly distributing the available capital over all available assets (the so-called $\frac{1}{n}$-portfolio), see also also [62]. In [37], Giannone et al. argue that the reason for this is the aforementioned instability. Therefore, the $\frac{1}{n}$-portfolio is considered a tough benchmark.

To test their model they use historical stock market data compiled by Fama and French [24]. They solve (5.2) using an ℓ^1-penalised regression method (*homotopy/LARS algorithm*, see [37, Appendix]). For this optimization the expected return is estimated by $\hat{\mu} = \frac{1}{T}\sum_{t=1}^T r_t$ over a sample period $1,\dots,T$. The desired return ρ of the portfolio is set to the average return of the $\frac{1}{n}$-portfolio over this period. Using the model (5.2) the authors construct portfolios using a sample period of 5 years. After the portfolio is constructed its return over the following year is computed from the set of historical data. This construction process is repeated at different times and the authors obtain a time series containing the returns of the portfolios. As a measure for the performance of the portfolios they compute the *Sharpe ratio* of the returns in this time series. The Sharpe ratio is given by the ratio between average mean return and the standard deviation.

The Sharpe ratio is then compared to the Sharpe ratio of the evenly distributed $\frac{1}{n}$-portfolio. Their results show that the portfolios obtained by the ℓ^1-approach outperform the $\frac{1}{n}$-portfolio for various time periods and different sets of possible assets. For a comparison of further regularisation approaches, and their computational results for the data sets by Fama and French, see [56].

We use the complementarity formulation of cardinality constrained optimization problems to obtain sparse portfolios. To evaluate the performance of the portfolios obtained by this approach we also use the Fama and French data set. We restrict the maximum number of active positions to $\kappa < n$ and consider the complementarity reformulation given by

$$\begin{aligned}
\min_{x,y} \ \|\rho e_T - Rx\|_2^2 \quad \text{s.t.} \quad & e^T x = 1, \\
& \mu^T x = \rho, \\
& x \geq 0, \\
& e^T y \geq n - \kappa, \\
& 0 \leq y_i \leq 1, \qquad \forall i = 1, \dots, n, \\
& x_i \cdot y_i = 0, \qquad \forall i = 1, \dots, n.
\end{aligned} \tag{5.3}$$

For this numerical study we add the constraint $x \geq 0$. This means that portfolio weights which include short selling are not feasible.

The optimization problems (5.2) and (5.3) represent two clearly different approaches. For (5.2) the number of active positions of the portfolio is not known a priori. Using (5.3) as a model the investor can fix the maximum number of active positions in the portfolio as a constraint. A further feature of the model (5.3) is the possibility to use *partial cardinality constraints*. In practice an investor might wish to choose a maximum number of active positions for different subsets of all available assets – for instance limiting the maximum number of assets from a certain industry, geographic region or traded in a given currency.

To evaluate the model, we use two data sets of historical returns compiled by Fama and French. We will compute portfolios for two selections of assets: The 48 Industry Portfolios (FF48) data set and the 100 Portfolios Formed on Size and Book-to-Market (FF100) data set.[1]

To evaluate the portfolios obtained by (5.3) we use the following scheme. For an evaluation period 1976 to 1986 using the FF48 data, for example, the starting point for the first portfolio construction is July 1976 using the historical data from July 1971 to June 1976. This means we are using a sample period of 5 years and, in the setting of (5.3), we have $n = 48$ and $T = 60$. The matrix R contains the historical returns for the sample period July 1971 to June 1976. As the desired return ρ we use the average monthly expected return of the $\frac{1}{n}$-portfolio over this time period given by

$$\rho = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{N} \sum_{i=1}^{n} r_{t,i}.$$

The desired portfolio weights in this model are then given by a solution of (5.3). To compute a solution we will use the Scholtes-type regularisation discussed in Section 4.3. Because the objective function is quadratic and the constraints are linear, the second order condition in Theorem 4.29 is satisfied. The gradients of the equality constraints are given by μ and e. Assuming $\mu \neq e$, CC-MFCQ likely holds at a local minimum x^*. Thus the requirements for the convergence of the Scholtes-type regularisation are most likely fulfilled, see Theorem 4.29.

After the portfolio is constructed, i.e. after a solution to (5.3) is computed, the annual return of this portfolio from July 1976 till June 1977 are recorded. The point in time for the next portfolio construction is then July 1977 and the process is repeated.

By proceeding in the same fashion for the whole period till July 1986, we compute a data series of annual returns. For our evaluation we compute the mean average, the standard

---

[1] Specifically, we use the sets *48 Industry Portfolios* (FF48) and *100 Portfolios Formed on Size and Book-to-Market* (FF100) (both *including* dividends). From these we work with the average value weighted monthly returns. The data sets are available at [24].

deviation and their ratio. These measures reflect the results that the investor would have achieved *on average* during the specific time span.
The described computation scheme was conducted for several time periods and for a range of parameters κ. The results for the data set FF48 are presented in Table 5.1 for values $\kappa = 4, 6, 8, 10, 12$. The portfolio construction was conducted for the time period 1976 to 2016 as well as for the subperiods 1976 to 1986, 1986 to 1996, 1996 to 2006 and 2006 to 2016. We compare the returns of the portfolios with returns of the equal weights $\frac{1}{n}$-portfolio. Each table presents the mean average m, the standard deviation σ and the Sharpe ratio m/σ of the portfolios' returns.
According to the same procedure, sparse portfolios were constructed for the data set FF100. The results are presented in Table 5.2.
For the data set FF48 the sparse portfolio achieves a higher Sharpe ratio for a majority of the time spans and constraints on the active positions. While the mean value of the $\frac{1}{n}$-portfolio's returns is higher in most cases, the sparse portfolios returns exhibit a lower standard deviation, resulting in a higher Sharpe ratio. For the data set FF100 one can observe a similar behaviour. For this data set the sparse portfolio's returns reach higher mean values while having a lower standard deviation than the $\frac{1}{n}$-portfolio without exception.
The portfolios, obtained by using the complementarity formulation as a model, perform well compared to the evenly distributed portfolio. Since the $\frac{1}{n}$-portfolio is a challenging benchmark, see [22], the results indicate the favourable performance of sparse portfolios and further support the findings in [37]. Moreover, the results indicate that the complementarity formulation serves well as a model for sparse portfolios.

	$\frac{1}{n}$-Portfolio			Sparse Portfolios														
				$\kappa = 4$			$\kappa = 6$			$\kappa = 8$			$\kappa = 10$			$\kappa = 12$		
Time span	m	σ	m/σ	m	σ	m/σ	m	σ	m/σ	m	σ	m/σ	m	σ	m/σ	m	σ	m/σ
1976 – 1986	1.553	4.731	0.328	1.324	3.532	0.375	1.436	3.533	0.406	1.430	3.512	0.407	1.425	3.511	0.406	1.427	3.505	0.407
1986 – 1996	1.286	4.489	0.287	1.088	3.470	0.314	1.072	3.385	0.317	1.112	3.371	0.330	1.104	3.358	0.329	1.111	3.360	0.331
1996 – 2006	1.038	4.069	0.255	0.735	3.999	0.184	0.772	3.510	0.220	0.746	3.492	0.214	0.788	3.422	0.230	0.777	3.422	0.227
2006 – 2016	0.886	4.995	0.177	0.752	3.647	0.206	0.937	3.497	0.268	0.940	3.420	0.275	0.962	3.411	0.282	0.963	3.405	0.283
1976 – 2016	1.099	4.631	0.237	0.914	3.680	0.248	1.013	3.506	0.289	1.010	3.474	0.291	1.025	3.449	0.297	1.025	3.447	0.297

Table 5.1: The table presents measures for annual returns of two portfolios for the **FF48** data set. The sparse portfolio with κ active positions and no short sales is obtained by computing a solution of (5.3). In the $\frac{1}{n}$-portfolio all positions are active with equal weights. The annual returns were computed for several time spans and different values of κ. For these returns the table presents mean average m, the standard deviation σ as well as the Sharpe ratio m/σ.

	$\frac{1}{n}$-Portfolio			Sparse Portfolios														
				$\kappa = 4$			$\kappa = 6$			$\kappa = 8$			$\kappa = 10$			$\kappa = 12$		
Time span	m	σ	m/σ	m	σ	m/σ	m	σ	m/σ	m	σ	m/σ	m	σ	m/σ	m	σ	m/σ
1976 – 1986	1.673	4.754	0.352	1.353	3.849	0.351	1.414	3.935	0.359	1.418	3.928	0.361	1.420	3.922	0.362	1.422	3.918	0.363
1986 – 1996	1.269	4.500	0.282	1.300	4.157	0.313	1.341	4.150	0.323	1.353	4.131	0.327	1.360	4.136	0.329	1.349	4.135	0.326
1996 – 2006	0.882	4.812	0.183	1.389	4.140	0.336	1.080	4.104	0.263	1.106	4.144	0.267	1.091	4.137	0.264	1.091	4.136	0.264
2006 – 2016	0.666	5.338	0.125	0.621	4.285	0.145	0.655	4.293	0.153	0.661	4.261	0.155	0.660	4.244	0.156	0.660	4.246	0.155
1976 – 2016	1.075	4.946	0.217	1.100	4.121	0.267	1.038	4.145	0.251	1.049	4.144	0.253	1.047	4.137	0.253	1.045	4.136	0.253

Table 5.2: The table presents measures for annual returns of two portfolios for the **FF100** data set. The sparse portfolio with κ active positions and no short sales is obtained by computing a solution of (5.3). In the $\frac{1}{n}$-portfolio all positions are active with equal weights. The annual returns were computed for several time spans and different values of κ. For these returns the table presents mean average m, the standard deviation σ as well as the Sharpe ratio m/σ.

5.2 Comparison of Methods for the Complementarity Formulation

In this section we will present results for the custom numerical methods for the complementarity formulation from Chapter 4. In Chapter 3 we discussed that standard constraint qualifications cannot be expected to hold for the complementarity formulation. These standard constraint qualifications are typically prerequisites for the convergence of numerical methods for nonlinear programs. We therefore studied custom constraint qualifications and stationarity conditions and, in Chapter 4, considered numerical methods which are based on these conditions. In particular we proved convergence for methods such as the ℓ^1-penalty method and regularisation methods under CC-constraint qualifications.
With the following numerical study we like to complement the theoretical discussion in the previous chapters. We investigate if the theoretical properties we discussed are reflected in the numerical results. The following custom numerical methods for the complementarity formulation are considered: The distance-based penalty formulation and the ℓ^1-penalty formulation from Section 4.1 as well as the Scholtes-type regularisation, the Kanzow-Schwartz regularisation and the exponential regularisation from Section 4.3.
We will apply these methods to optimization problems for sparse portfolio selection. For the portfolio optimization problems we are considering a similar study was conducted in [11].
In these models the variable x represents investments in n given assets, of which the expected return is modelled by a random variable. The expected return is given by $\mu \in \mathbb{R}^n$ and the covariance matrix by $Q \in \mathbb{R}^{n\times n}$. The objective function of the following model depends μ and Q. In practice both are estimated from historical data. An investor aims to achieve a high expected return while minimising a risk measure for the investment. The portfolio selection problem is modelled by the following cardinality constrained optimization problem:

$$\begin{aligned} \min_x \; r_\beta(x) \quad \text{s.t.} \quad & e^T x = 1, \\ & 0 \le x \le u, \\ & \|x\|_0 \le \kappa. \end{aligned} \tag{5.4}$$

The investor aims to minimise a risk measure of the portfolio given by r_β (depending on a parameter β), while spending the entire (normalised) budget which is modelled by the first constraint. We additionally assume an upper bound for x given by $u \in \mathbb{R}^n$ as well as non negativity of x, i.e. no short-sales. The cardinality constraint ensures that a portfolio has at most κ active positions. We reformulate the cardinality constraint using the complementarity formulation and test the methods for (1.2) on the following problem:

$$\begin{aligned} \min_{x,y} \; r_\beta(x) \quad \text{s.t.} \quad & e^T x = 1, \\ & 0 \le x \le u, \\ & e^T y \ge n - \kappa, \\ & 0 \le y \le 1, \\ & x \circ y = 0. \end{aligned} \tag{5.5}$$

As risk measure r_β we will use value-at-risk, conditional value-at-risk as well as robust counterparts. Let the random variable $\xi \in \mathbb{R}^n$ represent the return rates of the assets. Then for portfolio weights $x \in \mathbb{R}^n$ the loss is given by

$$\omega(x, \xi) = -x^T \xi.$$

Let ξ follow a probability distribution π. For a given level of confidence $\beta \in (0,1)$ value-at-risk (VaR) is defined as the minimum loss of a portfolio x not exceeded with probability of at least β, i.e.

$$\mathrm{VaR}_\beta(x) := \min\{z \,:\, P_\pi(\omega(x,\xi) \leq z) \geq \beta\}.$$

Typical values of β are close to 1. A further risk measure we are considering is conditional value-at-risk (CVaR), which is defined as the expected value of loss exceeding $\mathrm{VaR}_\beta(x)$.
Assuming standard normal distribution of the returns, the risk measures VaR and CVaR be can given in a closed form, see [11, Section 4] and the references therein. Let ϕ be the distributions density and Φ be the cumulative distribution function of the standard normal distribution. For the level of confidence we assume $\beta > 0.5$. Then we have

$$\mathrm{VaR}_\beta(x) = \zeta_\beta \sqrt{x^T Q x} - \mu^T x,$$

with $\zeta_\beta = -\Phi^{-1}(1-\beta)$. For conditional value-at-risk we have

$$\mathrm{CVaR}_\beta(x) = \eta_\beta \sqrt{x^T Q x} - \mu^T x,$$

where $\eta_\beta = (1-\beta)^{-1} \int_{-\infty}^{\Phi^{-1}(1-\beta)} t\phi(t)dt$.
Additionally to VaR and CVaR we are considering robust (or worst case) value-at-risk (RVaR) and robust (or worst case) conditional value-at-risk (CVaR), defined as

$$\begin{aligned}
\mathrm{RVaR}_\beta(x) &= \sup_\pi \mathrm{VaR}_\beta(x), \\
\mathrm{RCVaR}_\beta(x) &= \sup_\pi \mathrm{CVaR}_\beta(x).
\end{aligned}$$

Under the above assumptions on the underlying distribution, the following representations hold:

$$\begin{aligned}
\mathrm{RVaR}_\beta(x) &= \frac{2\beta - 1}{2\sqrt{\beta(1-\beta)}} \sqrt{x^T Q x} - \mu^T x, \\
\mathrm{RCVaR}_\beta(x) &= \sqrt{\frac{\beta}{1-\beta}} \sqrt{x^T Q x} - \mu^T x.
\end{aligned}$$

Using these four risk measures for different values for the level of confidence β we can generate a selection of test problems. As data for the covariance matrices and expected returns we use 30 randomly generated examples, also used in [33, 14, 11], which are available for the problem sizes $n \in \{200, 300, 400\}$ online at [32]. We test the methods on all three problem sizes and chose $\kappa = 10$ for the cardinality constraint. For the level of confidence we used the values $\beta \in \{0.9, 0.95, 0.99\}$ for each of the risk measures VaR, CVaR, RVaR and RCVaR. We therefore have 12 different objective functions for each of the 30 examples from [32], which results in 360 test cases.
For our study we will follow three possible approaches to compute a sparse portfolio. One possible approach is to regard (5.5) as a (standard) nonlinear optimization problem and apply a solver for nonlinear programs. Another approach is to use custom methods, the penalty and regularisation methods from Chapter 4, to compute a solution for (5.5).
We can follow these approaches for any optimization problem of the form (1.2) with continuously differentiable functions f, g and h, in particular including problems with nonlinear

functions. Problem (5.4) can be reformulated into a mixed-integer problem. Thus, in this case, we can additionally consider the mixed-integer formulation of (5.4) and apply a solver for mixed-integer programs.

The tested methods were implemented using the Python programming language together with its Scipy (version 0.16.0) and Numpy (version 1.8) libraries[2]. To solve the penalised problems (4.3) and (4.5), as well as the regularised problems occurring in the regularisation methods, we used the NLP Solver Snopt (version 7), see [38, 39]. We also used Snopt for the direct application of a NLP solver to (5.5). For the mixed-integer formulation we used the solver Gurobi (version 7.0.2), see [41]. All computations were run on the same computer (Intel Xeon CPU E5-2687W, 3.10GHz with 2x8 cores and 256GB memory).

5.2.1 Penalty Methods

We begin with the application of the ℓ^1-penalty approach and the distance-based penalty approach from Section 4.1 to (5.5). We applied the penalty methods as follows: We used the Solver Snopt to compute a solution of the penalised programs (4.3) respectively (4.5). As long as the computed solution was not feasible for the portfolio problem, we updated the penalty parameter according to $\alpha_{k+1} := 100.0 \cdot \alpha_k$. We started with the value $\alpha_0 = 100.0$. As start point for Snopt we used the last iterate. A point was accepted as feasible, if it satisfied all constraints up to tolerance of 10^{-6}. In particular the cardinality constraint was regarded as fulfilled, if $\min\{|x_i|, y_i\} \leq 10^{-6}$. For the penalty term in (4.3) we used the Euclidean norm. For all three approaches we started with the vectors $(0,0) \in \mathbb{R}^n \times \mathbb{R}^n$ and $(0,e) \in \mathbb{R}^n \times \mathbb{R}^n$.

As the problem instances are randomly generated we use aggregated attributes to compare the results. Tables 5.3, 5.4 and 5.5 present the average computation time, the number of times a given method computed the smallest objective function values, and the number of times an infeasible point was computed. Although all computations were made on the same computer, the computation times should be regarded as a qualitative measure only. Using the starting vector $x^0 = 0$ and $y^0 = 1$ for the penalty approach (4.3) did not result in feasible solutions. Therefore we omit it in the presentation.

Starting at $(x^0, y^0) = (0, e)$, the ℓ^1-penalty method computed the best solution compared to Snopt in circa 41% of the examples for = 200, 48% for $n = 300$ and 58% for $n = 400$. However, the best solutions for the majority of test cases was found by Snopt started with $(x^0, y^0) = (0, 0)$. The results indicate that the performance of the methods strongly depend on the starting vector. Regarding feasibility of the computed solutions, Snopt and the (distance-based) min-penalty approach delivered feasible points in all cases. Points computed with the ℓ^1 penalty approach were feasible in almost all cases as well.

To give a graphic comparison, we also use performance plots. Performance plots were introduced in [23] by Dolan and Moré. For each example j we identified the best objective function value f^j_{min} computed by any of the methods. For each method M and each example j we then computed the ratio f^j_M / f^j_{min}, where f^j_M is the objective value computed by method M for example j. The graph in the performance plot then reports the ratio of examples (on the vertical axis) for which this ratio is equal or lower to the value on the horizontal axis. If for a given example j a method does not compute a feasible point we set $f^j_M = +\infty$. Thus in case a method does not solve all problems, the performance plot does not reach the upper bound of 1.

[2]Scipy is a set of tools for scientific computing in Python (see `https://www.scipy.org`).

Thus the performance plots illustrate how good the solutions of a method are as well as how robust it is. We favour methods whose graph is close to the upper left corner, which means they deliver good and feasible solutions.
We present a performance plot of the objective function for the penalty methods and Snopt for the problem size $n = 400$ in Figure 5.1. Again, the dependence of the objective values on the starting vectors is noticeable. Although the ℓ^1-penalty approach with starting vector $(x^0, y^0) = (0, 0)$ delivers larger objective function values than Snopt in most cases, the plot shows that the gap is not too large. This also holds true for the solutions computed with the min-penalty approach.
To give a better comparison between the penalty methods we present aggregated results in Table 5.6, and a performance plot in Figure 5.2, which only contain the results for these methods. Although it did not compute the best solution in any of the examples, the min-penalty approach delivers good solutions in comparison to the ℓ^1-approach, which is indicated by the plot in Figure 5.2. Yet, the ℓ^1-penalty approaches seems to be superior: Table 5.6 shows that in all examples the ℓ^1-penalty approach computed the best solutions. The min-penalty approach with starting vector $(x^0, y^0) = (0, e)$ did not compute a feasible point for any of the examples.

	CVaR			RCVaR			RVaR			VaR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Penalty ℓ^1 (Start 0-0)	3.83	3.95	3.83	3.86	4.55	10.76	3.66	4.05	4.59	6.83	6.82	6.82
Penalty ℓ^1 (Start 0-1)	4.26	4.15	6.35	6.17	14.36	33.78	3.48	4.02	16.82	5.55	6.87	8.05
Penalty min (Start 0-0)	10.41	10.87	10.90	10.85	11.40	10.65	12.28	11.32	10.56	15.74	13.99	12.51
Snopt (Start 0-0)	0.98	0.96	1.02	1.22	1.15	1.04	1.12	1.12	1.10	2.43	2.53	2.37
Snopt (Start 0-1)	0.68	0.66	0.68	0.81	0.84	0.69	0.72	0.71	0.71	1.76	1.82	1.65
Algorithm	Best solution found											
Penalty ℓ^1 (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Penalty ℓ^1 (Start 0-1)	1	0	3	6	10	30	0	0	12	0	0	0
Penalty min (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-0)	29	30	27	24	20	0	30	30	18	30	30	30
Snopt (Start 0-1)	0	0	0	0	0	0	0	0	0	0	0	0
Algorithm	Computed point was infeasible											
Penalty ℓ^1 (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Penalty ℓ^1 (Start 0-1)	0	0	1	0	1	0	1	0	2	0	0	0
Penalty min (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-1)	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.3: Aggregated results of the penalty methods and Snopt for $n = 200$, $\kappa = 10$.

	CVAR			RCVAR			RVAR			VAR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Penalty ℓ^1 (Start 0-0)	12.92	15.79	12.80	19.01	17.41	54.08	13.41	13.80	23.73	17.30	16.83	17.85
Penalty ℓ^1 (Start 0-1)	17.60	23.75	29.05	42.06	65.18	141.10	13.90	22.96	68.61	17.50	21.75	32.49
Penalty min (Start 0-0)	33.78	31.59	31.21	31.75	31.25	38.78	34.91	34.54	36.88	36.40	35.31	34.37
Snopt (Start 0-0)	2.38	2.51	2.34	2.64	2.56	2.52	2.40	2.57	2.56	4.53	4.42	4.45
Snopt (Start 0-1)	1.84	1.86	1.77	1.91	1.74	1.69	1.98	1.90	1.68	3.54	3.33	3.35
Algorithm	Best solution found											
Penalty ℓ^1 (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Penalty ℓ^1 (Start 0-1)	0	0	10	12	19	30	0	0	22	0	0	5
Penalty min (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-0)	30	30	20	18	11	0	30	30	8	30	30	25
Snopt (Start 0-1)	0	0	0	0	0	0	0	0	0	0	0	0
Algorithm	Computed point was infeasible											
Penalty ℓ^1 (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Penalty ℓ^1 (Start 0-1)	0	2	0	0	0	0	0	2	0	1	1	0
Penalty min (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-1)	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.4: Aggregated results of the penalty methods and Snopt for $n = 300$, $\kappa = 10$.

	CVAR			RCVAR			RVAR			VAR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Penalty ℓ^1 (Start 0-0)	45.55	45.84	43.37	55.82	73.70	180.36	51.23	45.81	83.02	47.36	55.07	46.24
Penalty ℓ^1 (Start 0-1)	81.11	105.71	108.62	150.05	175.41	512.90	47.67	96.44	304.67	52.22	75.62	103.02
Penalty min (Start 0-0)	89.95	85.69	108.80	94.98	96.11	76.51	94.58	122.33	110.12	100.90	97.39	102.76
Snopt (Start 0-0)	4.97	5.22	4.94	5.13	5.14	5.25	4.91	4.99	5.27	7.53	7.80	7.57
Snopt (Start 0-1)	3.58	3.51	3.42	3.36	3.35	3.31	3.66	3.55	3.28	5.43	5.51	5.34
Algorithm	Best solution found											
Penalty ℓ^1 (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Penalty ℓ^1 (Start 0-1)	0	5	10	11	24	30	0	7	29	0	0	11
Penalty min (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-0)	30	25	20	19	6	0	30	23	1	30	30	19
Snopt (Start 0-1)	0	0	0	0	0	0	0	0	0	0	0	0
Algorithm	Computed point was infeasible											
Penalty ℓ^1 (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Penalty ℓ^1 (Start 0-1)	0	0	1	0	0	0	0	0	0	0	1	0
Penalty min (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Snopt (Start 0-1)	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.5: Aggregated results of the penalty methods and Snopt for $n = 400$, $\kappa = 10$.

	CVaR			RCVaR			RVaR			VaR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Penalty ℓ^1 (Start 0-0)	45.55	45.84	43.37	55.82	73.70	180.36	51.23	45.81	83.02	47.36	55.07	46.24
Penalty ℓ^1 (Start 0-1)	81.11	105.71	108.62	150.05	175.41	512.90	47.67	96.44	304.67	52.22	75.62	103.02
Penalty min (Start 0-0)	89.95	85.69	108.80	94.98	96.11	76.51	94.58	122.33	110.12	100.90	97.39	102.76
Algorithm	Best solution found											
Penalty ℓ^1 (Start 0-0)	30	25	20	19	6	0	30	23	1	30	30	19
Penalty ℓ^1 (Start 0-1)	0	5	10	11	24	30	0	7	29	0	0	11
Penalty min (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Algorithm	Computed point was infeasible											
Penalty ℓ^1 (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0
Penalty ℓ^1 (Start 0-1)	0	0	1	0	0	0	0	0	0	0	1	0
Penalty min (Start 0-0)	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.6: Aggregated results of the penalty methods for $n = 400$, $\kappa = 10$.

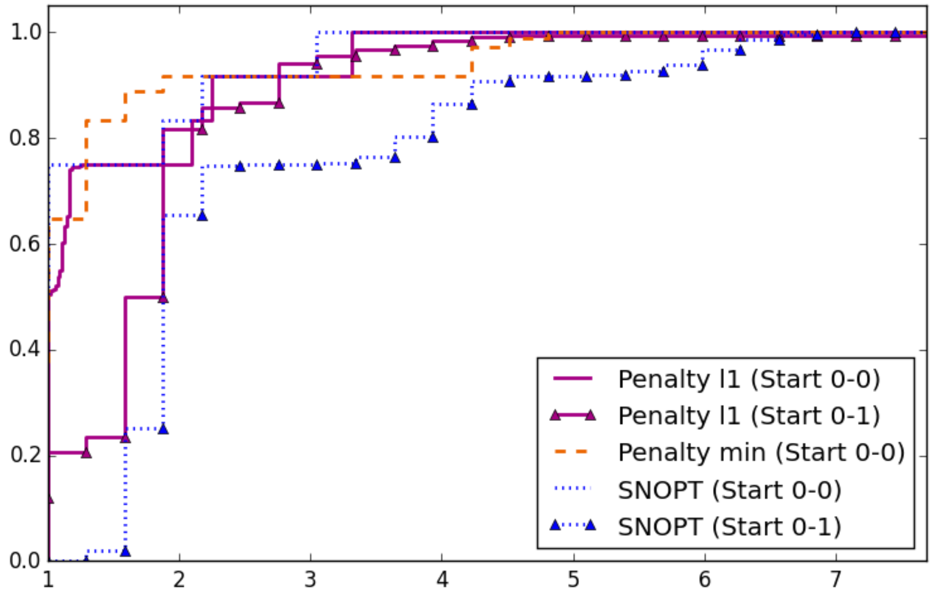


Figure 5.1: Performance plot of the objective function for the penalty methods and Snopt for $n = 400$, $\kappa = 10$.

5.2.2 Regularisation Methods

Regularisation methods for the complementarity formulation are a further class of methods we studied in Chapter 4. In this section we present results for the Scholtes-type regularisation, the Kanzow-Schwartz regularisation and the exponential regularisation. Again, we applied each method to each of the examples from [32] for every objective function VaR, CVaR, RVaR and RCVaR for three different values of β.

For the Scholtes-type regularisation and the Kanzow-Schwartz regularisation we proceeded as follows: In each iteration we applied the solver Snopt to the regularised problems NLP(t) for a regularisation parameter $t > 0$ starting with $t^0 = 1.0$. We accepted the computed point as feasible for (5.5), if it fulfilled all constraints up to a tolerance of 10^{-6}. For the complementarity constraint we checked if $\min\{|x_i|, y\} \leq 10^{-6}$ for all $i = 1, \ldots, n$. If the computed point was not feasible, we decreased the regularisation parameter according to $t^{k+1} = 0.01 \cdot t^k$ and again applied Snopt to the regularised problem. Increasing the parameter according to this rule yielded good results, see also the numerical results in [14, 15]. If no feasible point was computed after the parameter was decreased to 10^{-10}, we stopped.

For the exponential regularisation the parameter is *increased* if a computed solution for the regularised problems is infeasible. In this case we increased the parameter according to $t^{k+1} = \sigma \cdot t^k$ with $\sigma = 5$. If no feasible point was computed after the parameter was increased to 10^{10}, we stopped. A detailed study of this method can be found in [42] which includes a comparison of results for different values of σ. It was found that the exponential regularisation delivers better solutions for lower values of σ, i.e. if the regularisation parameter t is not increased too fast. The trade-off is that the computation time increases. In our study we could observe this behaviour as well. We choose $\sigma = 5$ to obtain computation times in the same magnitude of the computation times of the Scholtes-type regularisation and the

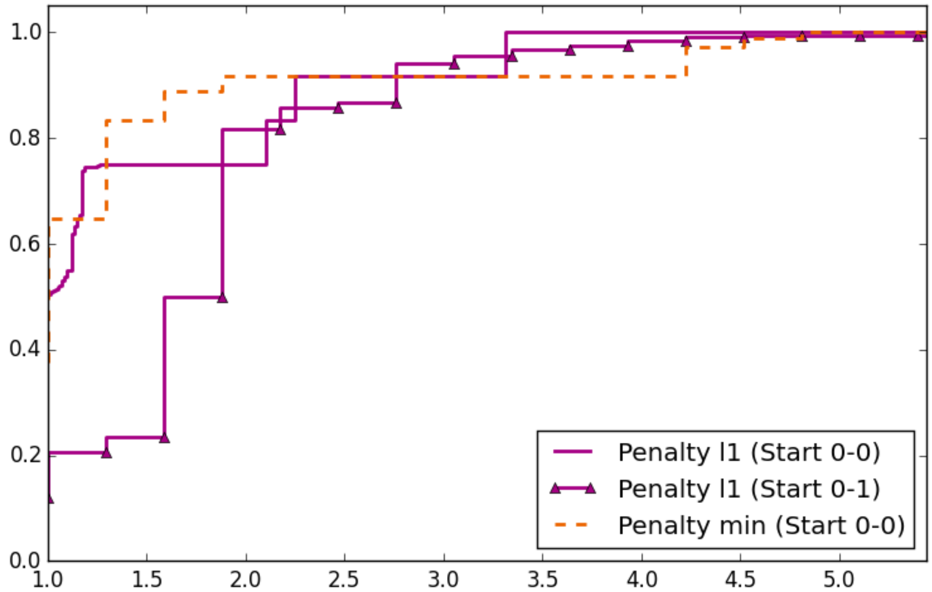


Figure 5.2: Performance plot of the objective function for the penalty methods for $n = 400$, $\kappa = 10$.

Kanzow-Schwartz regularisation.
As for the penalty methods, we used $(x^0, y^0) = (0, 0)$ and $(x^0, y^0) = (0, e)$ for all regularisation methods. Each of the methods worked best when started with $(x^0, y^0) = (0, 0)$.
As in the previous section, we present aggregated results in Tables 5.7, 5.8 and 5.9 and performance plots in Figures 5.3, 5.4 and 5.5 for the problem sizes $n \in \{200, 300, 400\}$ (for a explanation of performance plots see Section) 5.2.1.
For all examples and all objective functions every method successfully computed a feasible point for (5.5). The best solutions, for all problem sizes, were either computed by the Kanzow-Schwartz regularisation or by the Scholtes-type regularisation, which yielded the best solution of circa 60% of the examples. Regarding the computation times there are distinct differences. Snopt is by far the fastest method. Among the regularisation methods the Scholtes-type regularisation is fastest while the Kanzow-Schwartz regularisation the slowest. However, as mentioned before, for the exponential regularisation the computation time strongly depends on the increase of the regularisation parameter in each iteration. The performance plots in Figures 5.3, 5.4 and 5.5 show that, while it did not compute better solutions than the Scholtes-type and Kanzow-Schwartz regularisations, Snopt still delivered comparatively good solutions. For the problem size $n = 400$ the Scholtes-type regularisation, the Kanzow-Schwartz regularisation and Snopt behaved almost identical. This is also the case for the starting vector $(x^0, y^0) = (0, e)$ for which plots are contained in Figure 5.5.
The value $\kappa = 10$ for problem size $n = 200$ means 5% of the positions are allowed to be active. To evaluate the methods for different problem sizes $n \in \{200, 300, 400\}$ we also computed solutions to the test problems with $\kappa = 15$ for the problem size $n = 300$ and with $\kappa = 20$ for the problem size $n = 400$, hence keeping the ratio constant. For these cases, there was no qualitative difference in the performance of the Scholtes-type regularisation and Snopt.

	CVaR			RCVaR			RVaR			VaR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Reg. exponential	63.20	61.12	62.20	67.94	65.46	71.85	73.27	61.68	64.20	74.88	77.04	70.76
Reg. Kanzow-Schwartz	153.34	175.34	192.43	180.69	205.96	319.59	127.74	178.13	183.02	130.74	155.16	173.50
Reg. Scholtes	6.56	6.38	7.05	7.11	7.31	7.88	6.73	6.96	8.16	16.08	16.40	17.27
Snopt	0.98	0.96	1.02	1.22	1.15	1.04	1.12	1.12	1.10	2.43	2.53	2.37
Algorithm	Best solution found											
Reg. exponential	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Kanzow-Schwartz	11	9	10	12	12	12	13	12	11	11	10	12
Reg. Scholtes	19	21	20	18	18	18	17	18	19	19	20	18
Snopt	0	0	0	0	0	0	0	0	0	0	0	0
Algorithm	Computed point was infeasible											
Reg. exponential	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Kanzow-Schwartz	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Scholtes	0	0	0	0	0	0	0	0	0	0	0	0
Snopt	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.7: Aggregated results of the regularisation methods and Snopt for $n = 200$, $\kappa = 10$ and starting vector $x = 0$, $y = 0$

	CVaR			RCVaR			RVaR			VaR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Reg. exponential	174.72	173.50	188.90	196.03	152.14	206.47	152.53	182.05	150.03	169.17	192.72	193.42
Reg. Kanzow-Schwartz	414.57	462.33	436.22	433.78	479.59	627.43	327.96	451.57	508.37	325.33	431.38	438.00
Reg. Scholtes	18.18	18.85	19.57	19.26	18.47	19.80	16.82	18.89	19.78	30.64	34.70	34.94
Snopt	2.38	2.51	2.34	2.64	2.56	2.52	2.40	2.57	2.56	4.53	4.42	4.45
Algorithm	Best solution found											
Reg. exponential	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Kanzow-Schwartz	11	12	14	12	12	12	16	11	13	12	13	11
Reg. Scholtes	19	18	16	18	18	18	14	19	17	18	17	19
Snopt	0	0	0	0	0	0	0	0	0	0	0	0
Algorithm	Computed point was infeasible											
Reg. exponential	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Kanzow-Schwartz	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Scholtes	0	0	0	0	0	0	0	0	0	0	0	0
Snopt	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.8: Aggregated results of the regularisation methods and Snopt for $n = 300$, $\kappa = 10$ and starting vector $x = 0$, $y = 0$

	CVaR			RCVaR			RVaR			VaR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Reg. exponential	316.52	295.75	282.52	283.78	584.50	275.76	296.25	285.48	284.85	322.99	303.69	292.88
Reg. Kanzow-Schwartz	771.80	880.66	814.43	864.07	897.10	1207.09	704.02	812.32	900.34	691.79	838.90	856.38
Reg. Scholtes	38.03	38.24	39.94	38.75	39.23	42.64	36.83	38.40	41.39	57.92	59.16	61.14
Snopt	4.97	5.22	4.94	5.13	5.14	5.25	4.91	4.99	5.27	7.53	7.80	7.57
Algorithm	Best solution found											
Reg. exponential	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Kanzow-Schwartz	19	17	17	17	16	15	18	17	16	17	17	16
Reg. Scholtes	11	13	13	13	14	15	12	13	14	13	13	14
Snopt	0	0	0	0	0	0	0	0	0	0	0	0
Algorithm	Computed point was infeasible											
Reg. exponential	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Kanzow-Schwartz	0	0	0	0	0	0	0	0	0	0	0	0
Reg. Scholtes	0	0	0	0	0	0	0	0	0	0	0	0
Snopt	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.9: Aggregated results of the regularisation methods and Snopt for $n = 400$, $\kappa = 10$ and starting vector $x = 0$, $y = 0$

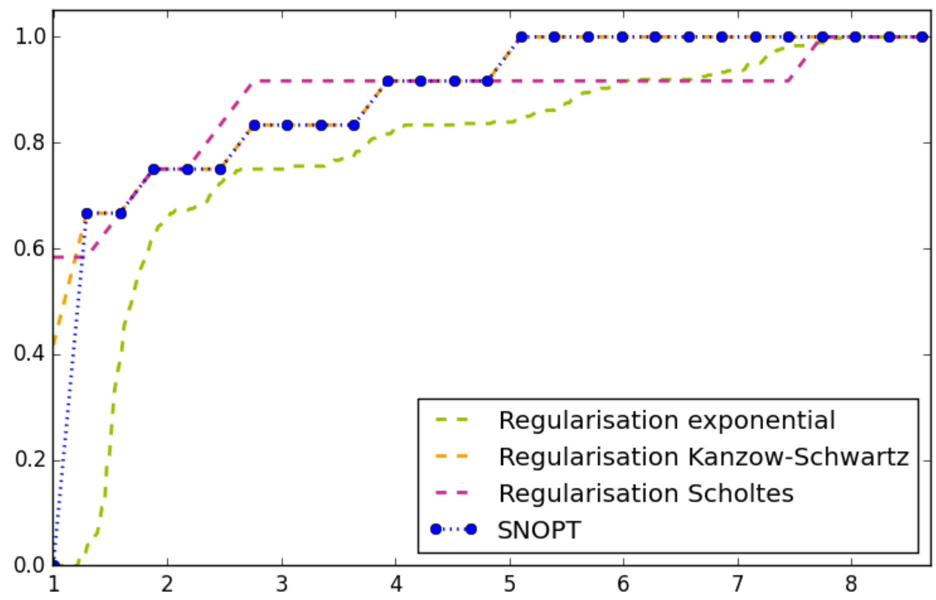


Figure 5.3: Performance plot of the objective function for the regularisation methods and Snopt for $n = 200$, $\kappa = 10$ with start vector $(0, 0)$.

A further possibility is to consider a mixed-integer formulation of (5.5). Since upper and lower bounds on x are given and the objective function is quadratic, we can use Gurobi for the following reformulation:

$$\begin{aligned} \min_{x,z,w,v} \quad c_\beta \cdot v - \mu^T x \quad \text{s.t.} \quad & e^T x = 1, \\ & 0 \le x \le u \circ z, \\ & e^T z \le \kappa, \\ & v \ge 0, \\ & w = Q^{\frac{1}{2}} x, \\ & v^2 \ge w^T w, \end{aligned} \tag{5.6}$$

where $x, w \in \mathbb{R}^n$, $v \in \mathbb{R}$ and the binary variable z is used to count the *nonzero* components of the vector x. The constant

$$c_\beta \in \left\{ \zeta_\beta,\, \eta_\beta,\, \frac{2\beta - 1}{2\sqrt{\beta(1-\beta)}},\, \sqrt{\frac{\beta}{1-\beta}} \right\}$$

depends on the respective risk measure. Gurobi is a *global* solver for mixed-integer programs which, if granted enough time, will find a global minimum. To obtain a baseline value for comparison with the *local* methods for the complementarity formulation, we are interested how Gurobi performs if granted roughly the same time. Therefore, we set Gurobi's timelimit option to circa the average computation time that the Scholtes-type regularisation required for each of the objective functions. To compute good solutions fast, we further set the MIPfocus option to 1. As starting vector for Gurobi we used $(x^0, y^0, w^0, v^0) = (0, e, 0, 0)$. Since z in (5.6) counts the nonzero components of x, and y in (5.5) counts the zero components, the start vector corresponds with the start vector $(x^0, y^0) = (0, 0)$ for the Scholtes-type regularisation.

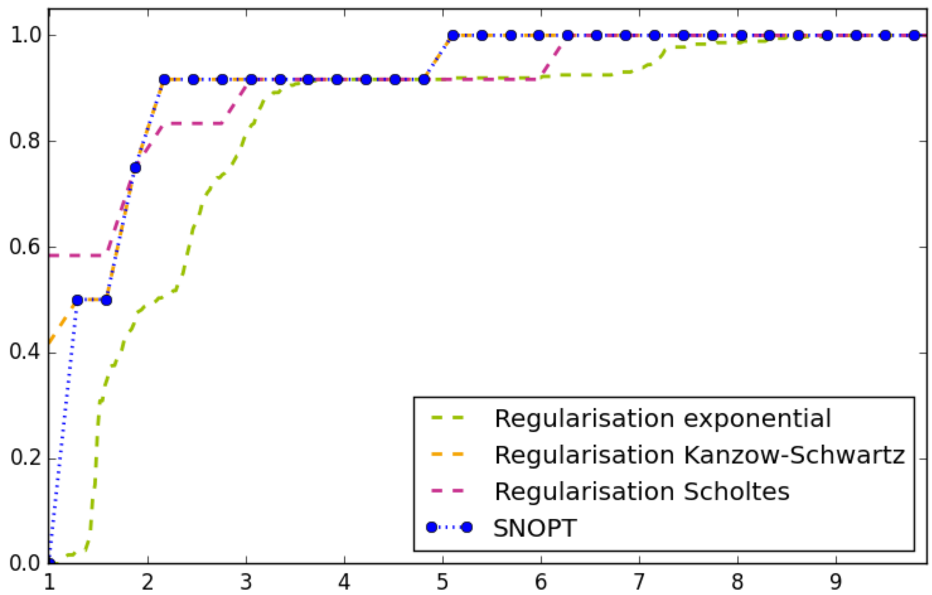


Figure 5.4: Performance plot of the objective function for the regularisation methods and Snopt for $n = 300$, $\kappa = 10$ with start vector $(0, 0)$.

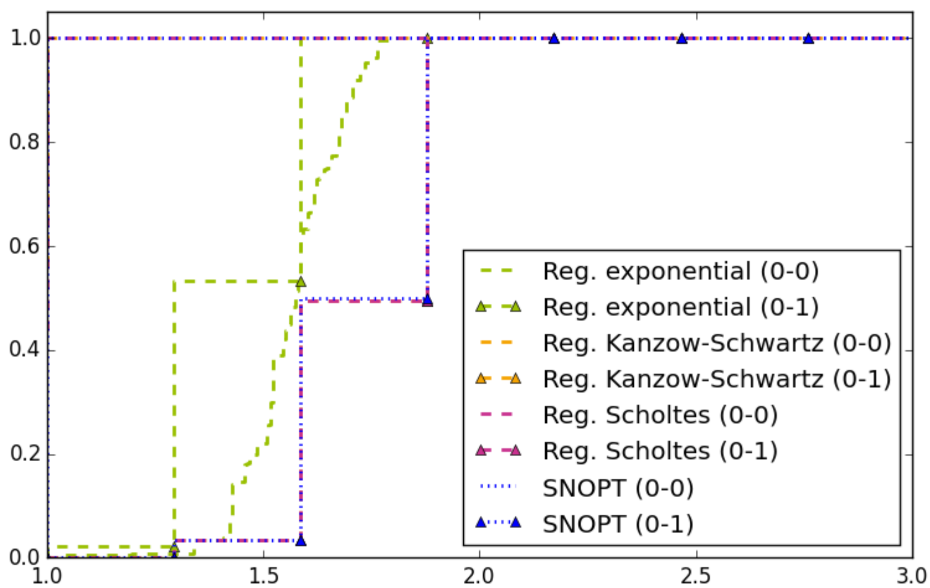


Figure 5.5: Performance plot of the objective function for the regularisation methods and Snopt for $n = 400$, $\kappa = 10$ with start vectors $(0, 0)$ and $(0, e)$.

	CVaR			RCVaR			RVaR			VaR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Gurobi	8.36	8.13	8.93	9.01	9.09	9.68	8.52	8.93	10.00	17.79	18.11	18.91
Reg. Scholtes	6.56	6.38	7.05	7.11	7.31	7.88	6.73	6.96	8.16	16.08	16.40	17.27
Algorithm	Best solution found											
Gurobi	4	9	5	8	8	6	6	8	8	8	7	6
Reg. Scholtes	26	21	25	22	22	24	24	22	22	22	23	24
Algorithm	Computed point was infeasible											
Gurobi	1	1	1	3	2	2	1	1	1	1	3	2
Reg. Scholtes	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.10: Aggregated results for the comparison between Gurobi and the Scholtes-type regularisation for $n = 200$, $\kappa = 10$.

	CVaR			RCVaR			RVaR			VaR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Gurobi	22.60	24.06	24.69	24.62	23.36	24.71	21.36	23.65	24.75	35.39	39.28	38.96
Reg. Scholtes	18.18	18.85	19.57	19.26	18.47	19.80	16.82	18.89	19.78	30.64	34.70	34.94
Algorithm	Best solution found											
Gurobi	4	8	0	4	4	7	8	3	5	6	5	13
Reg. Scholtes	26	22	30	26	26	23	22	27	25	24	25	17
Algorithm	Computed point was infeasible											
Gurobi	1	0	0	0	1	0	0	0	0	1	0	0
Reg. Scholtes	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.11: Aggregated results for the comparison between Gurobi and the Scholtes-type regularisation for $n = 300$, $\kappa = 10$.

	CVaR			RCVaR			RVaR			VaR		
	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Algorithm	Average computation time											
Gurobi	47.96	47.77	49.20	47.90	48.43	52.23	46.59	47.35	50.54	67.28	68.23	69.72
Reg. Scholtes	38.03	38.24	39.94	38.75	39.23	42.64	36.83	38.40	41.39	57.92	59.16	61.14
Algorithm	Best solution found											
Gurobi	5	6	6	9	7	4	5	8	5	8	7	7
Reg. Scholtes	25	24	24	21	23	26	25	22	25	22	23	23
Algorithm	Computed point was infeasible											
Gurobi	0	2	1	1	2	3	0	1	2	1	2	1
Reg. Scholtes	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.12: Aggregated results for the comparison between Gurobi and the Scholtes-type regularisation for $n = 400$, $\kappa = 10$.

Tables 5.10, 5.11 and 5.12 contain the aggregated results for this comparison of Gurobi and the Scholtes-type regularisation. The results for the Scholtes-type regularisation are from the previous comparison of the regularisation methods with Snopt, i.e. they were computed according to the same procedure. The tables show that, if granted circa the same amount of time, the Scholtes-type regularisation compares well to Gurobi regarding the number of examples in which the best solution is found. Figures 5.6, 5.7 and 5.8 contain performance plots for the results of Gurobi and the Scholtes-type regularisations for all problem sizes. The plots further indicate the good performance of the Scholtes-type regularisation.

In the same fashion, we further compared the Scholtes-type regularisation and Gurobi for all test problems with a value of 15 for κ for the problem size $n = 300$, and with a value of 20 for κ for the problem size $n = 400$, respectively, i.e. we kept the ratio of active positions at 5%. For this constant ratio between κ and n the Scholtes-type regularisation performed slightly better than Gurobi for both sizes $n = 300$ and $n = 400$.

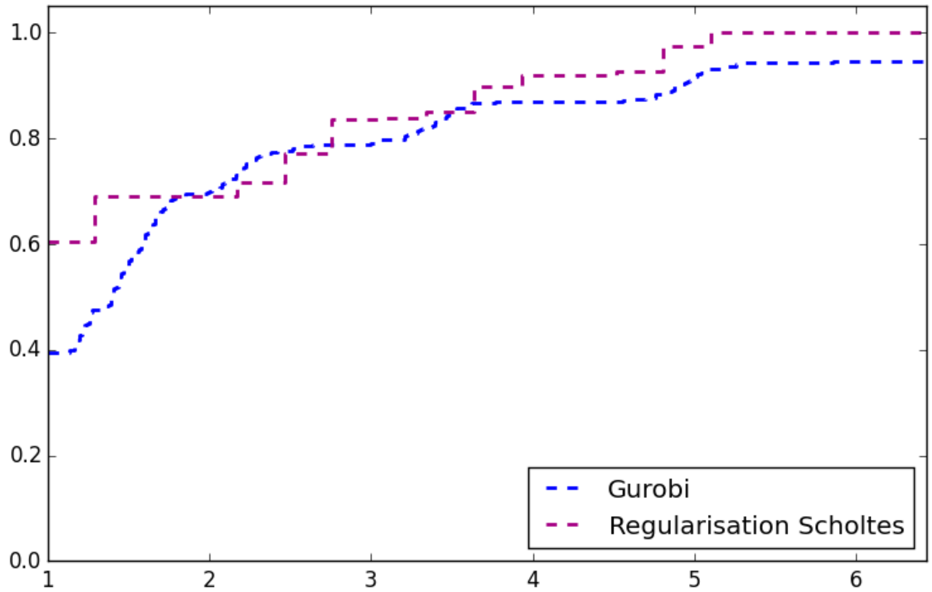


Figure 5.6: Performance plot of the objective function for the comparison of Gurobi and the Scholtes-type regularisation for $n = 200$, $\kappa = 10$.

Concluding Remarks on the Computational Results

In the first part of this chapter we used (1.2) as a model for sparse portfolio selection. Using historical stock market data we constructed portfolios that can compete with an evenly distributed portfolio, which is considered a tough benchmark, in terms of the Sharpe ratio. This comparison shows that the complementarity formulation serves well as a model for sparse portfolios.

In the second part of this chapter we considered a class of cardinality constrained portfolio optimization problems and applied different numerical methods to them. We used the penalty and regularisation methods discussed in Chapter 4 to compute solutions of the complementarity formulations of these problems. We further applied the nonlinear solver Snopt directly to

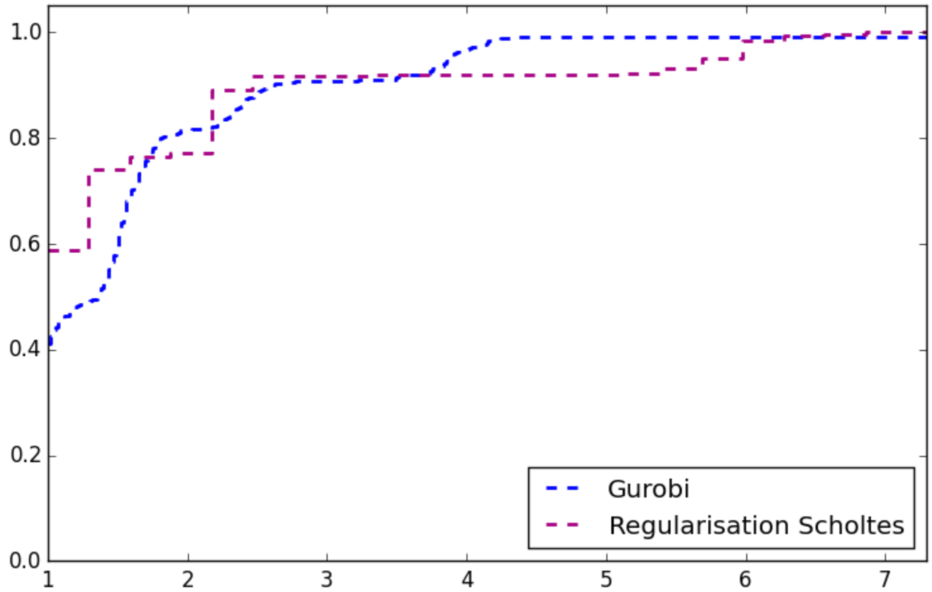


Figure 5.7: Performance plot of the objective function for the comparison of Gurobi and the Scholtes-type regularisation for $n = 300$, $\kappa = 10$.

the complementarity formulation as well as using the mixed-integer solver Gurobi to compute solutions of a mixed-integer formulation.
As to be expected for local methods, for the penalty methods, regularisation methods and Snopt the results are strongly dependent on the starting points. Starting at $(x^0, y^0) = (0, e)$, the ℓ^1-penalty method computed the best solution compared to Snopt in about half of the cases. However, Snopt starting at $(x^0, y^0) = (0, 0)$ yielded better solutions for a considerably larger number of examples. Among the penalty methods the ℓ^1-penalty approach worked best, especially regarding feasibility of the computed solutions. However, it was only able to outperform Snopt for particular objective functions.
The results of the regularisation methods show that their application to the complementarity formulation does work in practice. In particular the Kanzow-Schwartz regularisation and the Scholtes-type regularisation yielded good solutions in comparison to Snopt. Among the regularisation methods the Scholtes-type regularisation required the shortest computation times on average. This favourable numerical performance of the Scholtes-type regularisation corresponds to observations made for MPCCs [46]. Moreover, if the aim is to compute a good solution (but not necessarily the global solution) in a short amount of time, the Scholtes-type regularisation can compete with a commercial mixed-integer solver such as Gurobi.

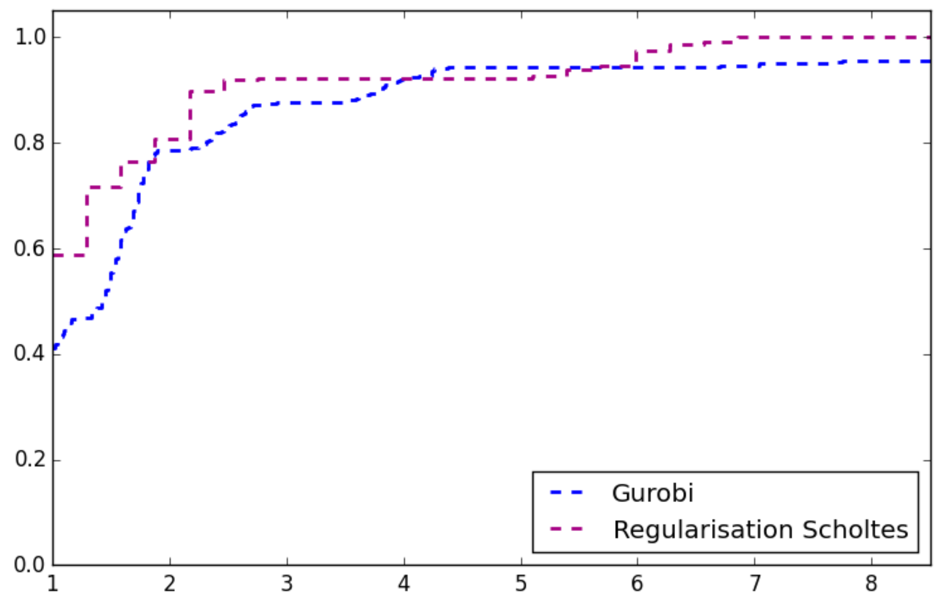


Figure 5.8: Performance plot of the objective function for the comparison of Gurobi and the Scholtes-type regularisation for $n = 400$, $\kappa = 10$.

6 Conclusion

The focus of this thesis was on a continuous reformulation of cardinality constrained optimization problems. This reformulation allows the application of numerical methods from nonlinear optimization. Therefore, nonlinear problems are included in our setting.
We successfully expanded the set of optimality conditions for the complementarity formulation. These new conditions include a second order necessary optimality condition, a second order sufficient optimality condition for S-stationary points and a uniqueness result for M-stationary points. Counterparts of these conditions with respect to the x variable only were derived as well. The optimality conditions rely on custom stationary conditions and hold under CC-constraint qualifications. Since we use a smaller critical cone than the corresponding result from the theory on standard nonlinear programs, these results hold under weaker conditions. In this way, we also take the lack of curvature of the objective function with respect to y into account. The second order sufficient optimality condition, for instance, can be used to check if an S-stationary point is in fact a local minimum. Additionally, we used second order conditions to derive convergence results for regularisation methods in the chapter on numerical methods. Moreover, regarding theoretical results, we proved a result on the existence of a local error bound for the complementarity formulation.
Furthermore, we considered numerical methods for the complementarity formulation. Using the derived local error bound we could prove exactness of a distance-base penalty function. For the special case $x \geq 0$ we considered an ℓ^1-penalty approach as well. This case is, for example, of interest in portfolio optimization problems. We showed that the limit of a sequence of KKT points of penalised problems is S-stationary, if it is feasible and CC-MFCQ holds there. This approach is similar to penalisation methods for MPCCs. In our case, we were able to obtain a stronger convergence result. We could show that, if the sequence converges to a feasible point of the complementarity formulation, this point is S-stationary. This result is stronger than in the MPCC case, where one can prove C-stationarity, and needs additional assumptions to show M-stationarity in the limit, see [48].
We further studied regularisation methods in detail. We considered a Scholtes-type regularisation which is known to perform well for MPCCs. For the Scholtes-type regularisation we proved convergence of KKT points of the regularised problems to an S-stationary point under CC-MFCQ. Using a similar proof, we also showed convergence of a regularisation that uses an exponential function to S-stationary points.
Using the second order optimality conditions we derived, we could expand the convergence theory of the Scholtes-type regularisation: We showed that the regularised problems posses a solution in the vicinity of a strict local solution of the cardinality constrained problem. Moreover, we proved that KKT points of the regularised problems exist and, under a second order condition, converge locally to this minimum. The same line of argument can in fact be used for a whole class of regularisation methods, which we specifically did for the Kanzow-Schwartz regularisation.
The convergence result presently known for the Kanzow-Schwartz regularisation only yields an M-stationary limit point. However, in contrast to the Scholtes-type regularisation it only

requires the weaker constraint qualification CC-CPLD to hold. If it is possible to show convergence to an S-stationary point under a stronger constraint qualification, such as CC-MFCQ, remains an open question. It is also not known if and how the convergence results of the discussed regularisations change, if one computes a sequence of approximations to KKT points. This could be a subject of future research.

We also considered the application of a (standard) SQP method to the complementarity formulation. Using a piecewise decomposition, we investigated its behaviour when applied to the complementarity formulation.

In the last chapter we conducted a numerical study. Firstly, we used the complementarity formulation as a model for sparse portfolio selection. Based on historical stock market data we constructed portfolios for a range of time spans. To evaluate their performance we compared their Sharpe ratios to the Sharpe ratio of an evenly distributed portfolio. The results confirm the good performance of sparse portfolios and show that the complementarity formulation serves well as a model for sparse portfolio selection.

Secondly, we compared numerical methods for the complementarity formulation. We considered the penalty and regularisation methods whose theoretical properties were studied in the chapter on numerical methods. We also applied Snopt, a solver for nonlinear programs, directly to the complementarity formulation. For a range of portfolio optimization test problems with different risk measures we compared the results of these methods. For particular objective functions the ℓ^1-penalty method delivered better results than Snopt. Although Snopt delivered better solutions in the majority of cases, penalty functions could be of further interest as merit functions for an SQP method. Regarding the regularisation methods, especially the Scholtes-type regularisation and the Kanzow-Schwartz regularisation compared well to Snopt. The Scholtes-type regularisation delivered the best solutions in the majority of cases while also requiring a short computation time on average. The good performance of this type of regularisation has been observed for MPCCs. In our case, however, this behaviour is also supported by a corresponding theoretical result.

Bibliography

[1] L. Adam and M. Branda. Nonlinear chance constrained problems: Optimality conditions, regularization and solvers. *Journal of Optimization Theory and Applications*, 170(2):419–436, 2016.

[2] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Mathematical Programming*, 135(1-2):255–273, 2012.

[3] R. Andreani, J. Martinez, and M. Schuverdt. On the relation between constant positive linear dependence condition and quasinormality constraint qualification. *Journal of Optimization Theory and Applications*, 125(2):473–483, 2005.

[4] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang. Restricted normal cones and sparsity optimization with affine constraints. *Foundations of Computational Mathematics*, 14(1):63–83, 2014.

[5] M. S. Bazaraa and C. M. Shetty. *Foundations of Optimization.* Lecture Notes in Economics and Mathematical Systems. Springer, 1976.

[6] A. Beck and Y. C. Eldar. Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *SIAM Journal on Optimization*, 23(3):1480–1509, 2013.

[7] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, Mass., 2. edition, 1999.

[8] D. P. Bertsekas and A. E. Ozdaglar. Pseudonormality and a Lagrange multiplier theory for constrained optimization. *Journal of Optimization Theory and Applications*, 114(2):287–343, 2002.

[9] D. Bertsimas and R. Shioda. Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1):1–22, 2009.

[10] D. Bienstock. Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, 74(2):121–140, 1996.

[11] M. Branda, M. Bucher, M. Červinka, and A. Schwartz. Convergence of a Scholtes-type regularization method for cardinality-constrained optimization problems with an application in sparse robust portfolio optimization. *Computational Optimization and Applications*, 70(2):503–530, 2018.

[12] M. Bucher, C. Kanzow, and A. Schwartz. Penalization techniques for a complementarity formulation of cardinality constrained optimization problems. in preparation, 2018.

[13] M. Bucher and A. Schwartz. Second-order optimality conditions and improved convergence results for regularization methods for cardinality-constrained optimization problems. *Journal of Optimization Theory and Applications*, 178(2):383–410, 2018.

[14] O. P. Burdakov, C. Kanzow, and A. Schwartz. Mathematical programs with cardinality constraints: Reformulation by complementarity-type conditions and a regularization method. *SIAM Journal on Optimization*, 26(1):397–425, 2016.

[15] E. Candes and M. Wakin. An Introduction To Compressive Sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[16] M. Cervinka, C. Kanzow, and A. Schwartz. Constraint qualifications and optimality conditions for optimization problems with cardinality constraints. *Mathematical Programming*, 160:353–377, 2016.

[17] T.-J. Chang, N. Meade, J. Beasley, and Y. Sharaiha. Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, 27(13):1271–1302, 2000.

[18] N. Chieu and G. Lee. A relaxed constant positive Linear dependence constraint qualification for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 158(1):11–32, 2013.

[19] F. H. Clarke. *Optimization and Nonsmooth Analysis*, volume 5. Siam, 1990.

[20] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, 2004.

[21] V. Demiguel, M. P. Friedlander, F. J. Nogales, and S. Scholtes. A two-sided relaxation scheme for mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 16(2):587–609, 2005.

[22] V. DeMiguel, L. Garlappi, and R. Uppal. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953, 2009.

[23] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, 2002.

[24] E. F. Fama and K. R. French. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Retrieved 19 January 2018.

[25] B. Fastrich, S. Paterlini, and P. Winker. Constructing optimal sparse portfolios using regularization methods. *Computational Management Science*, 12(3):417–434, 2015.

[26] M. Feng, J. E. Mitchell, J.-S. Pang, X. Shen, and A. Wächter. Complementarity formulations of l0-norm optimization problems. *Industrial Engineering and Management Sciences. Technical Report. Northwestern University, Evanston, IL, USA*, 2013.

[27] T. Fischer and M. E. Pfetsch. On the structure of linear programs with overlapping cardinality constraints. Technical report, Technical report, Available on Optimization Online, 2017.

[28] M. Flegel and C. Kanzow. Abadie-type constraint qualification for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 124(3):595–614, 2005.

[29] M. L. Flegel and C. Kanzow. On the Guignard constraint qualification for mathematical programs with equilibrium constraints. *Optimization*, 54(6):517–534, 2005.

[30] R. Fletcher and S. Leyffer. Solving mathematical programs with complementarity constraints as nonlinear programs. *Optimization Methods and Software*, 19:pp. 15–40, 2004.

[31] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes. Local convergence of SQP methods for mathematical qrograms with equilibrium constraints. *SIAM Journal on Optimization*, 17(1):259–286, 2006.

[32] A. Frangioni and C. Gentile. Mean-variance problem with minimum buy-in constraints, data and documentation. http://www.di.unipi.it/optimize/Data/MV.html. Retrieved 15 July 2015.

[33] A. Frangioni and C. Gentile. SDP diagonalizations and perspective cuts for a class of nonseparable MIQP. *Operations Research Letters*, 35(2):181–185, 2007.

[34] D. Gade and S. Küçükyavuz. Formulations for dynamic lot sizing with service levels. *Naval Research Logistics (NRL)*, 60(2):87–101, 2013.

[35] M. Galati. *Decomposition Methods for Integer Linear Programming.* PhD thesis, Lehigh University, 2010.

[36] C. Geiger and C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben.* Springer, Berlin [u.a.], 2002.

[37] D. Giannone, C. De Mol, J. Brodie, I. Daubechies, and I. Loris. Sparse and stable Markowitz portfolios. Working Paper Series 936, European Central Bank, 2008.

[38] P. Gill, W. Murray, M. Saunders, and E. Wong. *SNOPT 7.5 User's Manual.* CCoM Technical Report, Center for Computational Mathematics, University of California, San Diego, vol. 15-3 edition.

[39] P. E. Gill, W. Murray, and M. A. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAMR*, 47(1):99–1006, 2005.

[40] L. Guo, G.-H. Lin, and J. J. Ye. Second-order optimality conditions for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 158(1):33–64, 2013.

[41] Gurobi Optimization Inc. *Gurobi Optimizer Reference Manual*, 2016.

[42] J. Herzog. Ein exponentielles Relaxierungsverfahren für kradinalitätsrestringierte Optimierungsprobleme. Master's thesis, Fachbereich Mathematik Technische Universität Darmstadt, Darmstadt, Germany, 2018.

[43] T. Hoheisel and C. Kanzow. First-and second-order optimality conditions for mathematical programs with vanishing constraints. *Applications of Mathematics*, 52(6):495–514, 2007.

[44] T. Hoheisel and C. Kanzow. On the Abadie and Guignard constraint qualifications for Mathematical Programmes with Vanishing Constraints. *Optimization*, 58(4):431–448, 2009.

[45] T. Hoheisel, C. Kanzow, and J. V. Outrata. Exact penalty results for mathematical programs with vanishing constraints. *Nonlinear Analysis: Theory, Methods & Applications*, 72(5):2514–2526, 2010.

[46] T. Hoheisel, C. Kanzow, and A. Schwartz. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. *Mathematical Programming*, 137(1):257–288, 2013.

[47] J. Hu. *On Linear Programs with Linear Complementarity Constraints.* PhD thesis, Rensselaer Polytechnic Institute, Troy, New York, 2008.

[48] X. M. Hu and D. Ralph. Convergence of a penalty method for mathematical programming with complementarity constraints. *Journal of Optimization Theory and Applications*, 123(2):365–390, 2004.

[49] A. F. Izmailov and M. V. Solodov. On error bounds and Newton-type methods for generalized Nash equilibrium problems. *Computational Optimization and Applications*, 59(1-2):201–218, 2014.

[50] R. Janin. *Directional derivative of the marginal function in nonlinear programming*, pages 110–126. Springer Berlin Heidelberg, Berlin, Heidelberg, 1984.

[51] H. Jiang and D. Ralph. Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. *SIAM Journal on Optimization*, 10(3):779–808, 2000.

[52] A. Kadrani, J.-P. Dussault, and A. Benchakroun. A new regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 20(1):78–103, 2009.

[53] C. Kanzow and A. Schwartz. Mathematical programs with equilibrium constraints. Enhanced Fritz John conditions, new constraint qualifications and improved exact penalty results. *SIAM Journal on Optimization*, 20:2730–2753, 2010.

[54] C. Kanzow and A. Schwartz. A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. *SIAM Journal on Optimization*, 23(2):770–798, 2013.

[55] C. Kanzow and A. Schwartz. The price of inexactness: Convergence properties of relaxation methods for mathematical programs with complementarity constraints revisited. *Mathematics of Operations Research*, 40(2):253–275, 2015.

[56] G. Kroker. Stabilität optimaler Portfolios. Master's thesis, Fachbereich Mathematik Technische Universität Darmstadt, Darmstadt, Germany, 2017.

[57] G.-H. Lin and M. Fukushima. A modified relaxation scheme for mathematical programs with complementarity constraints. *Annals of Operations Research*, 133(1):63–84, 2005.

[58] D. D. Lorenzo, G. Liuzzi, F. Rinaldi, F. Schoen, and M. Sciandrone. A concave optimization-based approach for sparse portfolio selection. *Optimization Methods and Software*, 27(6):983–1000, 2012.

[59] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Wquilibrium Constraints*. Cambridge Univ. Press, Cambridge [u.a.], 1996.

[60] O. L. Mangasarian. *Nonlinear Programming*. McGraw-Hill series in systems science. McGraw-Hill, New York, 1969.

[61] H. M. Markowitz. Porfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[62] R. Michaud. The Markowitz optimization enigma: Is optimized optimal? *Financial Analysts Journal*, 45(1):31–42, 1989.

[63] A. Miller. *Subset Selection in Regression*. Chapman & Hall/CRC Press, 2nd ed. edition, 2002.

[64] L. Minchenko and A. Tarakanov. On error bounds for quasinormal programs. *Journal of Optimization Theory and Applications*, 148(3):571–579, 2011.

[65] B. S. Morduchovič. *Variational Analysis and Generalized Differentiation : 1. Basic Theory*, volume 330 of *Die Grundlehren der mathematischen Wissenschaften*. Springer, Berlin [u.a.], 2006.

[66] W. Murray and H. Shek. A local relaxation method for the cardinality constrained portfolio optimization problem. *Computational Optimization and Applications*, 53(3):681–709, 2012.

[67] V. Nicoletta, E. Lanzarone, V. Bélanger, and A. Ruiz. A cardinality-constrained robust approach for the ambulance location and dispatching problem. In P. Cappanera, J. Li, A. Matta, E. Sahin, N. J. Vandaele, and F. Visintin, editors, *Health Care Systems Engineering*, pages 99–109, Cham, 2017. Springer International Publishing.

[68] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer series in operations research and financial engineering. Springer, New York [u.a.], 2. ed. edition, 2006.

[69] J. Outrata, M. Kočvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Nonconvex Optimization and its Applications. Kluwer Academic Publishers, 1998.

[70] F. Pan and S. Kasiviswanathan. Efficient placement of directional antennas in infrastructure-based wireless networks. In *Military Communications Conference, 2011 - MILCOM 2011*, pages 1796–1801, 2011.

[71] L. Pan, N. Xiu, and J. Fan. Optimality conditions for sparse nonlinear programming. *Science China Mathematics*, pages 1–18, 2017.

[72] L.-L. Pan, N.-H. Xiu, and S.-L. Zhou. On solutions of sparsity constrained optimization. *Journal of the Operations Research Society of China*, 3(4):421–439, 2015.

[73] J.-S. Pang. Error bounds in mathematical programming. *Mathematical Programming*, 79(1-3):299–332, 1997.

[74] J.-S. Pang and M. Fukushima. Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints. *Computational Optimization and Applications*, 13(1):111–136, 1999.

[75] L. Qi and Z. Wei. On the constant positive linear dependence condition and its application to SQP methods. *SIAM Journal on Optimization*, 10(4):963–981, 2000.

[76] D. Ralph and S. J. Wright. Some properties of regularization and penalization schemes for MPECs. *Optimization Methods and Software*, 19(5):527–556, 2004.

[77] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*, volume 317 of *Die Grundlehren der mathematischen Wissenschaften*. Springer, Berlin [u.a.], corr. 3. print. edition, 2009.

[78] R. Ruiz-Torrubiano, S. García-Moratilla, and A. Suárez. Optimization problems with cardinality constraints. In *Computational Intelligence in Optimization*, pages 105–130. Springer, 2010.

[79] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1), 2000.

[80] S. Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918–936, 2001.

[81] A. Schwartz. *Mathematical Programs with Complementarity Constraints : Theory, Methods and Applications*. PhD thesis, Universität Würzburg, Würzburg, Germany, 2011.

[82] S. Steffensen and M. Ulbrich. A new relaxation scheme for mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 20(5):2504–2539, 2010.

[83] C. Sun, R. Dai, and M. Mesbahi. Weighted network design with cardinality constraints via alternating direction method of multipliers. *IEEE Transactions on Control of Network Systems*, PP(99):1–1, 2018.

[84] X. Sun, X. Zheng, and D. Li. Recent advances in mathematical programming with Semi-continuous Variables and Cardinality Constraint. *Journal of the Operations Research Society of China*, 1(1):55–77, 2013.

[85] M. Ulbrich and S. Ulbrich. *Nichtlineare Optimierung*. Mathematik Kompakt. Birkhäuser, Basel, 2012.

[86] J. Weston, A. Elisseeff, B. Schölkopf, and P. Kaelbling. The use of zero-norm with linear models and kernel methods. *Journal of Machine Learning Research*, pages 1439–1461, 2003.

[87] J. J. Ye. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *Journal of Mathematical Analysis and Applications*, 307(1):350–369, 2005.

[88] G. Yuan and B. Ghanem. Sparsity constrained minimization via mathematical programming with equilibrium constraints. *arXiv preprint arXiv:1608.04430*, 2016.

[89] X. Zheng, X. Sun, D. Li, and J. Sun. Successive convex approximations to cardinality-constrained convex programs: a piecewise-linear DC approach. *Computational Optimization and Applications*, 59(1-2):379–397, 2014.

Wissenschaftlicher Werdegang

06/2018	Doctor rerum naturalium, Fachbereich Mathematik, Technische Universität Darmstadt
02/2018 - 06/2018	Wissenschaftlicher Mitarbeiter im Sonderforschungsbereich Transregio 154 der Deutschen Forschungsgemeinschaft (DFG): Mathematische Modellierung, Simulation und Optimierung am Beispiel von Gasnetzwerken, Technische Universität Darmstadt
02/2015 - 01/2018	Stipendiat der DFG im Rahmen der Graduate School Computational Engineering (DFG Project Graduate School 233), Technische Universität Darmstadt
04/2011 - 08/2014	Master of Science, Wirtschaftsmathematik, Karlsruher Institut für Technologie
03/2012 - 07/2012	Auslandssemester im Rahmen des Erasmusprogramms an der Universidade do Minho, Guimarães, Portugal
10/2007 - 02/2011	Bachelor of Science, Technische Volkswirtschaftslehre, Karlsruher Institut für Technologie
10/2005 - 09/2006	Studium der Physik, Ruprecht-Karls-Universität Heidelberg
05/2005	Abitur an der Otto-Kühne-Schule Godesberg, Bonn